BMJ Open

EU-funded initiatives for Real World Evidence: descriptive analysis of their characteristics and relevance for regulatory decision making

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2018-021864
Article Type:	Research
Date Submitted by the Author:	24-Jan-2018
Complete List of Authors:	Plueschke, Kelly; Pharmacovigilance and Epidemiology McGettigan, Patricia; European Medicines Agency, Pharmacovigilance and Epidemiology Department; William Harvey Research Institute, Barts, and The London School of Medicine and Dentistry, Queen Mary University of London Pacurariu, Alexandra; European Medicines Agency, Pharmacovigilance and Epidemiology Department; Dutch Medicines Evaluation Board Kurz, Xavier; European Medicines Agency, Cave, Alison; European Medicines Agency, Pharmacovigilance and Epidemiology Department
Keywords:	EU-funded initiatives, Real World Evidence, data sources, methodologies, governance and analytical models, infrastructure

EU-funded initiatives for Real World Evidence: descriptive analysis of their characteristics and relevance for regulatory decision making

Kelly Plueschke[1], Patricia McGettigan[1, 2], Alexandra Pacurariu[1, 3], Xavier Kurz[1], Alison Cave[1]

[1] Pharmacovigilance and Epidemiology Department, Inspections Human Medicines Pharmacovigilance and Committees Division, European Medicines Agency (EMA), London E14 5EU, United Kingdom

[2] William Harvey Research Institute, Barts, and The London School of Medicine and Dentistry, Queen Mary University of London, London EC1M 6BQ

[3] Dutch Medicines Evaluation Board, 3531 AH Utrecht, The Netherlands

Corresponding author:

Kelly Plueschke, kelly.plueschke@ema.europa.eu, +44 (0)20 3660 8531

Word count:

2632 (excluding title page, abstract, article summary, references, list of figures' and table's titles, funding, conflicts of interest, disclaimer, acknowledgement)

Number of figures and tables:

6 figures and 1 table in the manuscript

5 figures and 3 tables in the supplementary file

Key words:

EU-funded initiatives, Real World Evidence, data sources, methodologies, governance models, analytical models, infrastructure, sustainability

ABSTRACT

Introduction

A review of European Union (EU)-funded initiatives linked to 'Real World Evidence' (RWE) was performed to determine whether their outputs could be used for the generation of real-world data able to support the European Medicines Agency (EMA)'s regulatory decision making on medicines.

Method

The initiatives were identified from publicly available websites. Their topics were categorised into 5 areas: 'Data source', 'Methodology', 'Governance model', 'Analytical model' and 'Infrastructure'. To assess their immediate relevance for medicines evaluation, their therapeutic areas were compared to the products recommended for EU approval in 2016 and those included in the EMA pharmaceutical business pipeline.

Results

Of 171 originally identified EU-funded initiatives, 65 were selected based on their primary and secondary objectives (35 'Data source' initiatives, 15 'Methodology', 10 'Governance model', 17 'Analytical model' and 25 'Infrastructure'). These received 734 million Euros of public funding. At the time of evaluation, the published outputs of the 40 completed initiatives did not always match their original objectives. Overall, publicly-available information was limited, data access not explicit, and initiatives' sustainability unclear. The topics matched 8 of 14 therapeutic areas of the products recommended for approval in 2016, and of the 15 therapeutic areas in the 2017-2019 pharmaceutical business pipeline. Haematology, gastroenterology or cardiovascular system were poorly represented in the initiatives.

Conclusions

This landscape of EU-funded initiatives linked to RWE as of 31st December 2016 highlighted that the immediate utilisation of their outputs to support regulatory decision making is limited, often due to insufficient available information and discrepancies between outputs and objectives. Furthermore the restricted sustainability of the initiatives impacts on their downstream utility. Multiple projects focussing on the same therapeutic areas increases the likelihood of duplication of both efforts and resources. These issues contribute to gaps in generating RWE for medicines and diminish returns on the public funds invested.

ARTICLE SUMMARY

Strengths and limitations of this study

- This is the first evaluation of EU-funded initiatives linked to 'Real World Evidence' (RWE) that looks at the potential for their outputs to support regulatory decision making on medicines;
- The analysis is based on review of the publicly available information provided by each initiative, but in some cases this was limited in detail and in quality;
- There were discrepancies between the initiatives' objectives and their outputs, and a mismatch with the therapeutic areas of drugs recently recommended for approval or in the EMA's business pipeline;
- Despite funding totalling 734 million Euros for the 65 initiatives evaluated, most had limited sustainability and were dependent on new support to continue beyond the funding period;
- The article proposes solutions to enable better streamlining, communication and sustainability of the outputs generated through EU funded initiatives.

INTRODUCTION

The clinical evidence collected for the marketing authorisation of new medicines traditionally comes from randomised clinical trials (RCTs) but it is recognised that RCT data have limitations including tightly-controlled conditions of clinical care, highly selected populations, and in some scenarios, small sample sizes [1]. As a result, their applicability to the safety and efficacy of medicines in post-authorisation use is unknown. There is therefore a need to supplement RCTs with other sources more representative of everyday 'real world' medical practice in order to provide additional insight on the benefit-risk balance.

According to the GetReal Glossary of Definitions of Common Terms [2], 'Real World Evidence' (RWE) derives from the analysis and/or synthesis of 'Real World Data' (RWD) that can either be primary data collected in a manner which reflects how interventions would be used in routine clinical practice, or secondary data derived from routinely collected data. The range of RWD is wide and sources include electronic healthcare records, patient/disease registries, hospital records and health insurance data/claims databases. The EU Network Strategy to 2020 identifies RWE as a key enabler to bring innovative products to patients with unmet medical needs and to support the safe and effective use of medicines [3].

Whilst the EMA already uses RWE sources in its evaluations, this is typically on an 'ad hoc' basis. There is therefore a need to systematically understand RWE outputs at an EU wide level in order to make best use of existing information and to identify areas where further efforts are needed. We therefore created an inventory of EU-funded initiatives linked to RWE and RWD. The objectives were: i) to identify the initiatives established in terms of RWD sources, relevant methodologies, governance models, analytical models and infrastructures created to facilitate RWD collection, transformation, sharing, and analysis; ii) to understand if and how these initiatives could be exploited to support regulatory decision making on medicines [4]; iii) to consider the initiatives strategically in terms of needs, gaps and opportunities for the generation of RWD related to the therapeutic areas of medicinal products in the pharmaceutical development pipeline.

METHODS

Selection (1st step)

A 3-step approach was followed (Figure 1), starting with an internet search performed by one reviewer (KP) to identify completed and ongoing RWE initiatives funded through the 6th and 7th Framework Programmes (FP6/FP7) [5, 6], Horizon 2020 (H2020) [7] (including the Innovative Medicines Initiative (IMI) [8]), other EU initiatives (e.g. European Research Council, Strengthening Collaboration for Operating Pharmacovigilance in Europe (SCOPE) Joint Action [9], HIMSS EUROPE GOVERNING COUNCIL [10]) and nationally-funded initiatives discovered during the searches. The search terms used were: registry/ies, real world evidence, clinical, electronic medical record, eHealth, big data, electronic health record, data linkage/link, paediatrics, pregnancy, geriatrics, hospital data, data source, unique patient identifier, coding terminology, governance model, common protocol, distributing data approach, pooling of data, and analytical model. The search cut-off date was 31st December 2016.

Initiatives that matched any of the keywords were included in an Excel tracking table (Section *1.1., Supplementary file).* Based on their website information, they were categorised according to whether they had developed potentially relevant 1) RWD sources, 2) methodologies, 3) governance models, 4) analytical models or 5) infrastructure that could help provide evidence to support regulatory decision making (Section *1.2.1., Supplementary file)*. Initiatives could fall into more than one category. Those that did not focus on one or more of the categories or that focused only on the early stage of drug development (e.g. pre-clinical or clinical trials, development of molecular compounds/biomarkers), on the '-omics' (e.g. genomics), or on diseases with limited geographical spread (e.g. Ebola outbreak)

(Figure 2) were considered out of scope of the inventory and were therefore not selected for further analysis.

Category attributes (2nd step)

For each of the five categories, attributes were defined to facilitate selection of the outputs potentially most relevant for assisting in regulatory evaluations (for example for Data Sources, the attributes included collaborator access, database linkage, unique patent identifier, paediatrics data, and so on). Three reviewers (KP, PM, AP) with experience in pharmacovigilance and regulatory science identified the primary and secondary objectives of each initiative, the attributes, as well as the corresponding deliverables through examination of the initiatives' webpages and through reading of published documents including reports, presentations and publications. They then assigned the initiatives to one or more categories as appropriate *(*Sections *1.1* and *1.2.2., Supplementary file).* In cases of doubt, the three reviewers consulted to achieve consensus.

Scoring (3rd step)

A dashboard listing all the attributes in columns was created for each category. Each initiative was first categorised then characterised through identification and scoring of its attributes. Scores ranged from 0 to 2 for 'Yes / No' questions, for example, whether the data source included paediatric data (where 0='No', 1='Unknown' and 2='Yes'); or from 0 to 3 for more qualitative questions e.g. on the type of care covered (0='Unknown', 1='Primary', 2= 'Secondary/Tertiary' and 3='Primary, Secondary, Tertiary').

Two attributes were scored across all dashboards: 1) the availability of published outputs at the end of each completed project, and 2) whether the published outputs fully addressed the originally stated objectives. The dashboards permitted the creation of figures for each of the categories showing the number of initiatives fulfilling each attribute according to the information provided on their websites.

<Figure 1>

Initiative partners and funding

The numbers of partners named in each initiative and the amount of funding awarded were included in each dashboard.

Therapeutic Areas of the initiatives, new medicinal products and EMA business pipeline

The therapeutic areas of the initiatives included in the Data source category were mapped to:

- those of the products recommended for approval through the EU centralised authorisation procedure by the Committee for Medicinal Products for Human Use (CHMP) [11] in 2016;
- those of the products included in the EMA pharmaceutical business pipeline [12] through which developers have expressed a clear intent to submit a marketing authorisation application (MAA) through the centralised procedure to the EMA between March 2017 and December 2019.

RESULTS

The first screening returned 171 potentially relevant initiatives (Section *1.1., Supplementary file)* that matched at least one of the search keywords. Following the second screening, 115 initiatives were reviewed for categorisation (51 FP, 30 IMI, 15 H2020, 12 other EU initiatives, 7 nationally funded initiatives), of which 65 fell into one or more of the five categories (Figure 2). Some initiatives were included in more than one category, for example, data source initiatives that also developed infrastructure. The selected initiatives had a median duration of 5 years.

<Figure 2>

Dashboards scoring and summary of findings

Initiatives were scored in the dashboards according to their category(ies) attributes. The main findings for each category are summarised below and illustrated in Figure 3. The scoring dashboards are provided in Section *2.1.* of the *Supplementary file*.

- Data Sources: Information on access to the data sources developed through the initiative was available on the websites of 5 of the 35 Data source initiatives, including 2 where only the data guardian had access, and 3 where access was open to collaborators. Access information was either not publically available or unclear for the remaining 30 initiatives.
- Methodologies: Fifteen initiatives developed methodologies that could be applied to RWD studies, including guidance on protocol design (9), on the management of bias / confounders (5), and on the use of electronic health records and/or registry data (6). Seven initiatives referred to the use of established diseases or drug coding terminologies such as MedDRA, Snomed, ATC or Orphanet.
- Governance models: Ten initiatives dealt with confidentiality and data protection aspects, while 6 related to a new code of conduct.
- Analytical models: Ten of 17 initiatives aimed to develop smart text analysis tools such as machine learning, natural language programming, or data mining. Data transformation (anonymisation, right-protection and compression) was referred to in 9 initiatives.
- Infrastructure: All 25 initiatives related to the development of platforms or websites to share, extract and store data, of which 2 also mentioned cloud based technologies. Three developed a smartphone application that could be used by stakeholders including patients to record personal health information or report suspected adverse drug reactions.

<Figure 3>

Outputs versus objectives

The outputs matched the objectives for all 5 of the completed 'Governance model' initiatives (100%), 5 of the 6 completed 'Methodology' initiatives (83%), 10 of the 12 completed 'Infrastructure' initiatives (83%), 15 of the 23 completed 'Data source' initiatives (65%) and for 7 of the 12 completed 'Analytical model' initiatives (58%) (Table 1).

Table 1: Number of completed initiatives with outputs matching objectives as of 31st December 2016

Categories (number of initiatives)	Initiatives Completed / Ongoing	Output match objectives	Unclear if Output match objectives
Data source (n= 35)	23 / 12	15 (65%)	8 (35%)
Methodology (n=15)	6 / 9	5 (83%)	1 (17%)
Governance model (n=10)	5 / 5	5 (100%)	0
Analytical model (n=17)	12 / 5	7 (58%)	5 (42%)
Infrastructure (n=25)	12 / 13	10 (83%)	2 (17%)

Initiatives' sustainability

Whilst some initiatives continued beyond their agreed timelines through new EU-funded programmes within e.g. H2020, or were followed-up by existing networks (e.g. foundations, associations), others

ceased at the end of their allocated time and budget without a sustainability plan for any of the outputs delivered. Of the 23 completed 'Data source' initiatives, 7 had a clear sustainability mechanism described on their website, for example the EMIF Project [13]. Some more recent initiatives integrate roadmaps for future development and continuation as part of their deliverables, like RETHINK big [14].

Comparison of therapeutic areas

Of the 14 different therapeutic areas where products were recommended by the EMA for approval in 2016 and of the 15 represented in the EMA business pipeline, 8 were studied by the initiatives (Figures 4, 5). These included oncology (5 initiatives), neurology/central nervous system (CNS) (7), respiratory (5), immunology (5), rheumatology (3) and metabolic diseases (2). Other therapeutic areas where new products were about to be approved or marketing authorisation applications submitted were not represented in the initiatives, including in particular haematology, gastroenterology or cardiovascular system.

<Figures 4 and 5>

Number of partners and funding

Information on funding was available on the websites of 53 of the 65 initiatives categorised and scored. The total funding was 734 million Euros ranging from 1 to 56 million Euros. Information on the numbers of partners was provided for 58 initiatives and ranged from 1 to 161 partners.

For large initiatives (>20 partners), there was no clear relation between the number of partners and the amount of funding (Figure 6, funding of the 30 'Data source' initiatives). The same partners were involved in multiple initiatives, especially pharmaceutical companies as well as academic centres.

<Figure 6>

DISCUSSION

To our knowledge, this is the first evaluation of EU-funded initiatives in terms of their potential to develop RWD sources, methodologies, governance models, analytical models or infrastructure to contribute to the utility of RWE. It highlights discrepancies between the initiatives' objectives and outputs (as of 31st December 2016), a lack of sustainability of the outputs arising from the initiatives, and a mismatch with the therapeutic areas of drugs recently recommended for approval by the EMA and appearing in its business pipeline. Nevertheless the inventory can be consulted by EMA upon receipt of enquiries arising in the course of regulatory assessments of medicinal products; navigation through the 5 dashboards and filtering of the relevant attributes might provide sources of additional evidence to support decision-making.

Observation

Of 35 'Data source' initiatives, 27 developed RWD sources that focused on diseases in 8 specific therapeutic areas, more than the half of which (17) clustered in neurology/CNS, oncology, and respiratory (Figures 4 and 5). These multiple parallel initiatives with substantial number of partners sometimes involved in the same projects may increase the risk of 1) duplication of efforts and resources, 2) overlap of final outputs, and 3) the possibility of generating unsustainable and non-interoperable data sources.

The inventory highlights that some areas of new product development such as haematology, gastroenterology and cardiovascular system are not represented in the RWD sources identified in the inventory. Therefore opportunities to generate RWD to support evaluation and decision-making on new products could be created through new initiatives targeting these therapeutic areas. It is possible that

such omissions will be addressed to some extent by the strategic research agenda of the recently-launched IMI 2 [15] (e.g. Big Data for Better Outcomes Programme [16]).

Whilst the inventory focuses on EU funded initiatives, the same exercise could be extended to international initiatives in order to enlarge the geographical spread and number of data sources, provide a better understanding of populations studied in real world settings worldwide and widen the possibility of linkage between data sources through new analytical models.

Limitations

The three consecutive steps were based on the review of the websites by individual EMA staff members. However limited information published on the websites (e.g. lack of final reports or accounts of deliverables while some initiatives did not even have a website), broken links and both limited access / limited information on access to the data, made it sometimes difficult to determine the initiatives' attributes for inclusion in the inventory, let alone conclude on their general applicability to regulatory science. As the initiatives were publicly funded, it is reasonable to expect that their websites would provide up to date information including published outputs as well as some indication in relation to data access and sustainability matters. Adoption of the suggestion of Galsworthy et al. to develop a central EU repository to make outputs permanently accessible for open meta-analysis and data reuse would go some way to improving this deficit [17, 18]. For initiatives whose data contributes to post-authorisation studies, their registration in the publicly accessible EU PAS Register® provides an opportunity to increase the dissemination of methods and results [19].

Proposed way forward

Most recent initiatives include roadmaps as part of their deliverables and this will potentially assist in ensuring sustainability. Moreover the 11th Call for Proposals launched by the IMI 2 Joint Undertaking [20] aims to provide solutions to ensure that significant results from IMI projects become fully exploitable, available to all relevant end users, and/or fully sustainable in the long term and in their own right.

Based on our analysis, additional options to ensure better exploitation of these publically funded initiatives include the following:

- To streamline efforts, resources, and promote interoperability between outputs, consideration must be given to existing data and lessons learnt from past projects. This should help the identification of the real public health needs to be addressed by new initiatives.
- If timescales allow, future EU RWE initiatives could take account of the EMA business pipeline on medicinal products to avoid gaps in the generation of RWD and ensure regulatory needs are supported.
- The maintenance of the initiatives' websites at the end of the funding period is key to keep stakeholders up to date on the progress made and on the deliverables/achievements. Maintenance of these websites should permit easy public access to information, related reports and peer reviewed publications but would require dedicated funding.
- There is a need for clarity on the possibility to access RWD sources generated by the initiatives to allow their reuse in other contexts.
- Consideration of mechanisms to ensure the sustainability of outputs delivered by initiatives should be a priority and should be a requirement for all proposals.

CONCLUSION

The development of the inventory assists in understanding the extent of existing RWD resources emerging from EU funded initiatives and highlights multiple shortcomings. Ideally results of such initiatives should be reused and sustained and any lessons learnt disseminated. However despite the potential of some initiatives to provide RWE that would support decision-making on the safety of medicinal products, there are challenges for their utilisation in a regulatory context due to the obstacles in the exploitation of their outputs (e.g. limited data access, lack of sustainability).

Gaps and opportunities were identified in terms of specific therapeutic areas which may require RWD but were not a focus of funded initiatives. A number of solutions are proposed to enable better streamlining, communication and sustainability of the outputs generated through the EU funded initiatives. As the 65 initiatives together were granted 734 million Euros of public funding, it is imperative that the deficiencies highlighted here are addressed in future funding programs. This would go some way to ensuring delivery of stated objectives, data availability, sustainability and reflection of areas of medical need.

REFERENCES

1. Vandenbroucke, J.P., *Observational research, randomised trials, and two views of medical science.* PLoS Med, 2008. **5**(3): p. e67.
2. GetReal. *GetReal WP1 - Deliverable D1.3 "Glossary of Definitions of Common Terms"*. 2016; Available from: http://www.imi-getreal.eu/Portals/1/Documents/01%20deliverables/D1.3%20-%20Revised%20GetReal%20glossary%20-%20FINAL%20updated%20version_25Oct16_webversion.pdf.
3. *EU Medicines Agencies Network Strategy to 2020 - Working together to improve health*. December 2015; Available from: http://www.ema.europa.eu/docs/en_GB/document_library/Other/2015/12/WC500199060.pdf.
4. Goldman, M., N. Seigneuret, and H.G. Eichler, *The Innovative Medicines Initiative: an engine for regulatory science.* Nat Rev Drug Discov, 2015. **14**(1): p. 1-2.
5. Commission, E. *6th Framework Programmes (FP6)*. 2002-2006; Available from: https://ec.europa.eu/research/fp6/index_en.cfm.
6. Commission, E. *7th Framework Programmes (FP7)*. 2014-2020; Available from: https://ec.europa.eu/research/fp7/index_en.cfm.
7. Commission, E. *Horizon 2020 (H2020)*. 2018-2020; Available from: http://ec.europa.eu/programmes/horizon2020/.
8. *Innovative Medicines Initiative*. Available from: https://www.imi.europa.eu/.
9. SCOPE. *European Research Council, Strengthening Collaboration for Operating Pharmacovigilance in Europe* Available from: http://www.scopejointaction.eu/.
10. *HIMSS EUROPE GOVERNING COUNCIL*. Available from: http://www.himss.eu/about-himss/governing-council.
11. *Committee for Medicinal Products for Human Use (CHMP)*. Available from: http://www.ema.europa.eu/ema/index.jsp?curl=pages/about_us/general/general_content_000094.jsp&mid=WC0b01ac0580028c79.
12. *EMA publication on Business pipeline*. 2015; Available from: http://www.ema.europa.eu/docs/en_GB/document_library/Leaflet/2011/03/WC500104229.pdf .
13. *The EMIF Project*. Available from: http://www.emif.eu/.
14. *RETHINK Big*. Available from: http://www.rethinkbig-project.eu/.
15. *Strategic Research Agenda for Innovative Medicines Initiative 2*. Available from: http://www.imi.europa.eu/sites/default/files/uploads/documents/IMI2_SRA_March2014.pdf.
16. *Big Data for Better Outcomes Programme (BD4BO)*. Available from: http://bd4bo.eu/.
17. Galsworthy, M.J., et al., *Academic output of 9 years of EU investment into health research.* Lancet, 2012. **380**(9846): p. 971-2.
18. Molokhia, M., et al., *Academic output from EU-funded health research projects.* Lancet, 2012. **380**(9857): p. 1903-4.
19. Register, E.P. *The European Union electronic Register of Post-Authorisation Studies*. Available from: http://www.encepp.eu/encepp_studies/indexRegister.shtml.
20. *Innovative Medicines Initiatives 2 Joint Understanding 11th Call for Proposals on Exploitation of IMI project results*. Available from: https://www.imi.europa.eu/sites/default/files/uploads/documents/IMI2Call11/IMI2_Call11_CallText.pdf.

LIST OF FIGURES' AND TABLE'S TITLES

Figure 6 - Number of partners versus funding of 'Data source' initiatives

Figure 6 - Numerical data for the number of partners versus funding of Data source initiatives

FOOTNOTES

Acknowledgement: Kelly Plueschke thanks Enrico Tognana at the European Medicines Agency for providing data on the EMA pharmaceutical business pipeline covering the period from March 2017 and December 2019, as well as the peer-reviewers for the revision of this manuscript.

Contributions: AC and XK conceived the idea of the descriptive analysis. All authors were involved in the study design. KP, PM and AP performed the analysis and interpretation of results. All authors contributed to writing the manuscript and approved the final draft. KP is the guarantor.

Funding: No sources of funding were used to assist in the preparation of this article.

Conflict of interest: We have read and understood the BMJ Group policy on declaration of interests and declare the following interests: None

Disclaimer: The views expressed in this article are the personal views of the authors and may not be understood or quoted as being made on behalf of or reflecting the position of the European Medicines Agency or one of its committees or working parties.

Ethical approval: Not required.

Data sharing: All the information provided in this manuscript other than the data on the EMA business pipeline is publically available online.

Enquiries on the EMA business pipeline should be submitted to the EMA as Access-to-documents requests by completing the online form.

Transparency: KP affirms that the manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned (and, if relevant, registered) have been explained.

Patient involvement statement: This descriptive analysis did not involve any patients.

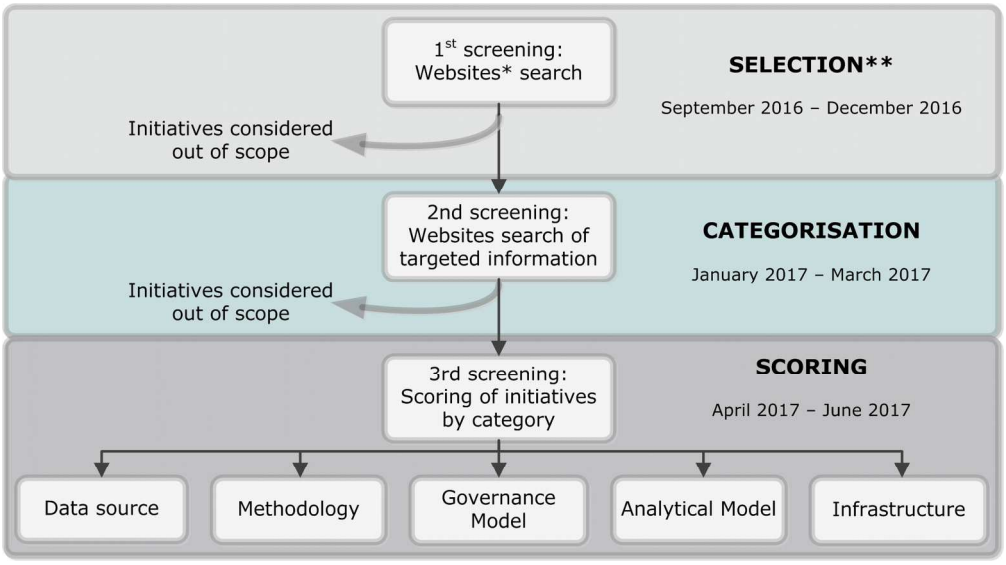

Figure 1 - Development steps of the inventory of EU-funded RWE initiatives

166x92mm (300 x 300 DPI)

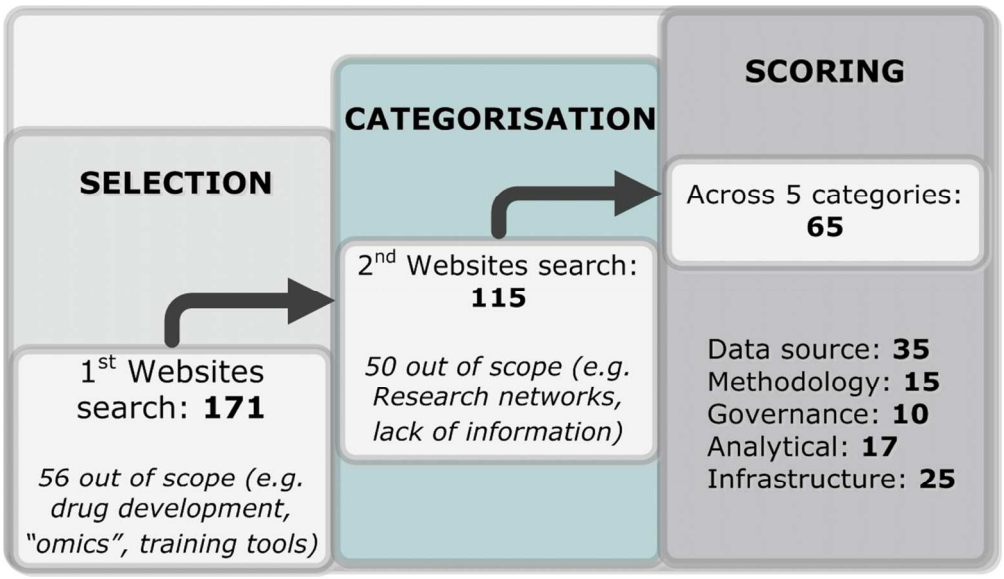

Figure 2 - Selections of initiatives for the inventory of EU-funded RWE initiatives

104x59mm (300 x 300 DPI)

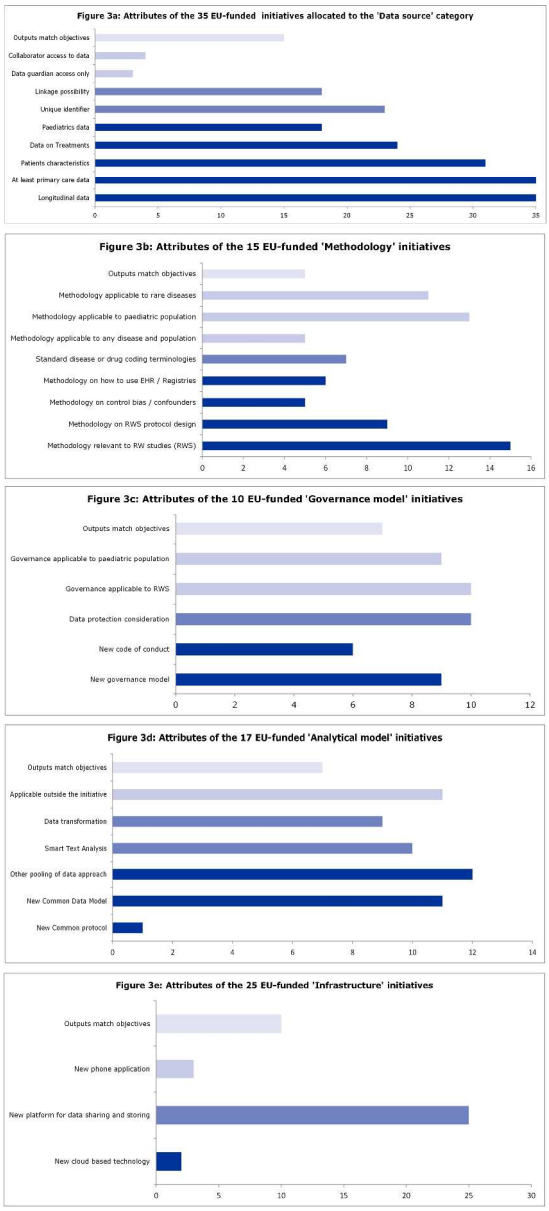

Figure 3 - Attributes of the EU-funded initiatives according to each of the 5 categories

219x484mm (300 x 300 DPI)

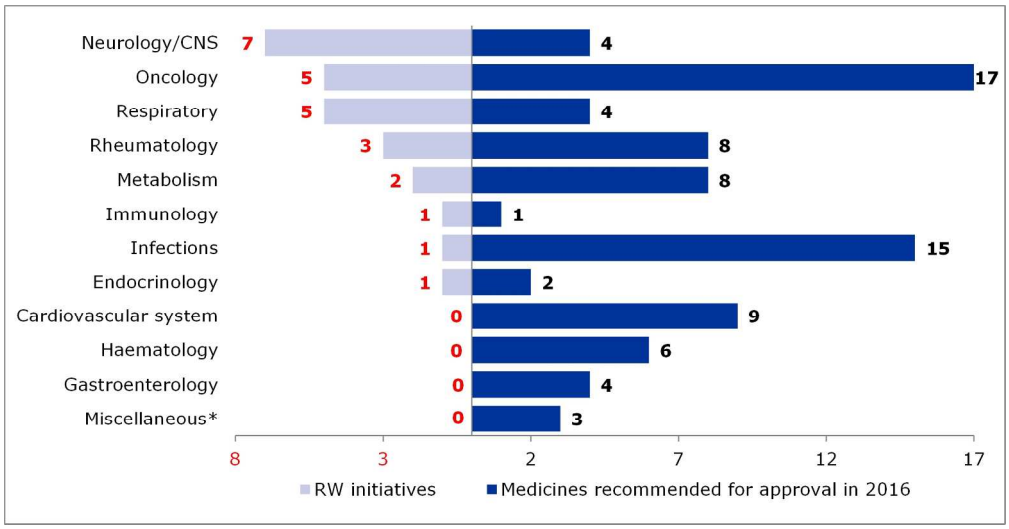

Figure 4 - Specific therapeutic areas of products recommended for approval in 2016 versus those covered by the identified initiatives

198x102mm (300 x 300 DPI)

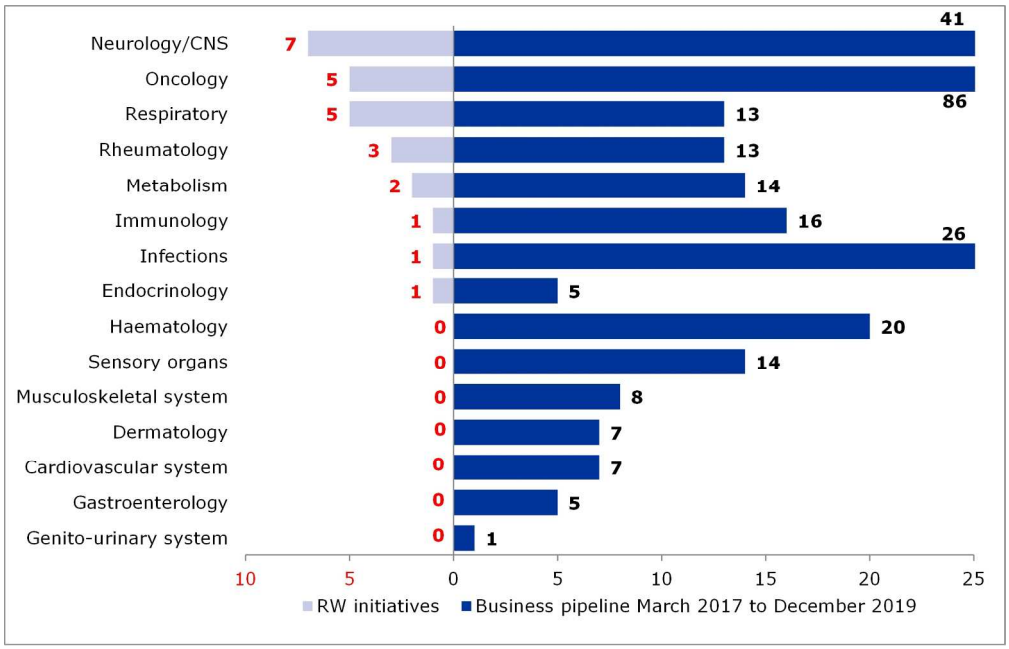

Figure 5 - Specific therapeutic areas of products included in the EMA business pipeline (March 2017 to December 2019) versus those covered by the identified initiatives

196x127mm (300 x 300 DPI)

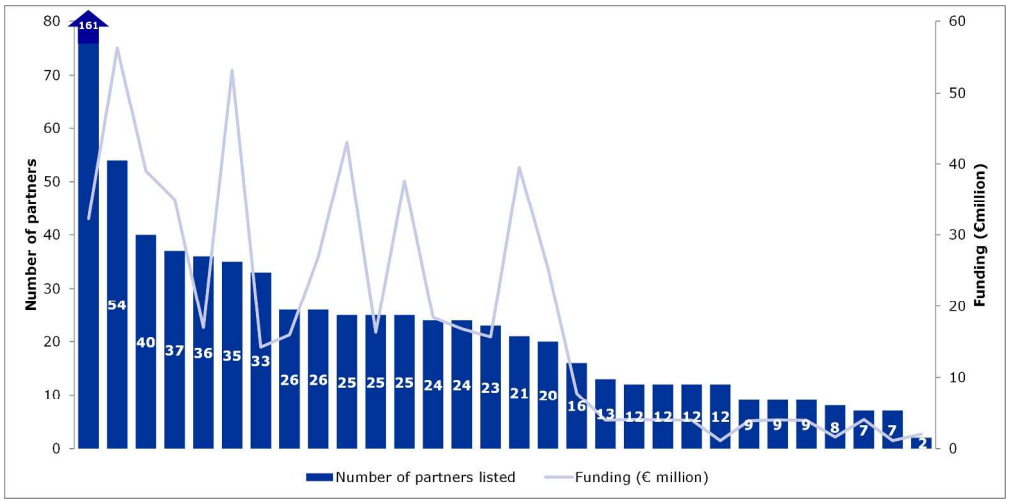

Figure 6 - Number of partners versus funding of Data source initiatives

252x124mm (300 x 300 DPI)

Supplementary file of the manuscript '*EU-funded initiatives for Real World Evidence: descriptive analysis of their characteristics and relevance for regulatory decision making'*

1. Information related to the Methodology section of the manuscript

1.1. List of EU-funded initiatives

The 171 initiatives which initially matched at least one of the keywords used in the internet search performed between September and December 2016 are listed in the table below, together with their EU-funding programme and a link to relevant websites or publications used during the analysis. Some of the initiatives were considered 'out of scope' of the inventory following the 'selection' step (56) or the 'categorisation' step (50). The remaining initiatives (65) underwent 'scoring' across the 5 categories, namely Category 1 'Data source', Category 2 'Methodology', Category 3 'Governance model', Category 4 'Analytical method' and Category 5 'Infrastructure'.

Initiative full name	Initiative acronym	Programme	Approximate start and end dates	Scoring (primary category in bold)
Anti-Biopharmaceutical Immunisation: Prediction and Analysis of Clinical Relevance to Minimise the Risk	ABIRISK	IMI	March 2012 – February 2017	Out of scope
Accelerated Development of Appropriate Patient Therapies -a Sustainable, Multi-stakeholder Approach from Research to Treatment-outcomes	ADAPT SMART	IMI	July 2015 – December 2017	Out of scope
Accelerated development of vaccine benefit-risk collaboration in Europe	ADVANCE	IMI	October 2013 - October 2018	3, 2, 5
Organising mechanistic knowledge about neurodegenerative diseases for the improvement of drug development and therapy	AETIONOMY	IMI	January 2014 - December 2019	Out of scope
Apoptosis systems biology applied to cancer and AIDS	APO-SYS	FP7	February 2008 – January 2012	Out of scope
Applied public-private research enabling osteoarthritis clinical headway	APPROACH	IMI	June 2015 – November 2020	1, 2, 5
Assessing SNOMED CT for Large Scale eHealth Deployments in the EU	ASSESS CT	H2020	February 2015 – December 2016	Out of scope
Advances in Small Trials dEsign for Regulatory Innovation	Asterix	FP7	January 2013 - September 2017	2

Initiative full name	Initiative acronym	Programme	Approximate start and end dates	Scoring (primary category in bold)
and eXcellence				
Assessment of the Safety of Labas in Asthma in Routine Care	Astrolab	FP7	January 2011 - May 2016	1
Big Data EEG Recording and Analysis platform	Big Data EEG – "WhiteBox EEG"	H2020	July 2016 – December 2016	Out of scope
Big data against childhood Obesity	BigO	H2020	January 2016 - November 2020	1, 2
Big Medical Data Use in Primary	BIMEDA	H2020	August 2015 - July 2017	Out of scope
BiobankCloud	BiobankCloud	FP7	December 2012 – November 2015	Out of scope
Building data bridges from biology to medicine in Europe	BIOMEDBRIDGES	FP7	January 2012 – December 2015	Out of scope
Biobank Standardisation and Harmonisation for Research Excellence in the European Union	BioSHaRE-EU	FP7	December 2010 – November 2015	Out of scope
Biomarkers for Enhanced Vaccine Immunosafety	BioVacSafe	IMI	March 2012 - February 2017	Out of scope
Epigenetic blueprint of haematopoietic cells	BLUEPRINT	FP7	October 2011 – September 2016	Out of scope
Be the Cure	BTCure	IMI	April 2011 - March 2017	1
C3-Cloud	C3-Cloud	H2020	May 2016 – April 2020	Out of scope
Collaborative European NeuroTrauma Effectiveness Research in Traumatic Brain Injury	CENTER-TBI	FP7	October 2013 - April 2020	1
Chemical manufacturing methods for the 21st century pharmaceutical industries	CHEM21	IMI	October 2012 – June 2017	Out of scope
Developing a Child Cohort Research Strategy for Europe	CHICOS	FP7	January 2010 - February 2013	1
European Management Platform for Childhood Interstitial Lung Diseases	chILD-EU	FP7	December 2012 - June 2016	1
Combatting Bacterial Resistance in Europe	COMBACTE	IMI	January 2013 – December 2019	Out of scope
Combatting Bacterial Resistance in Europe - Carbapenem Resistance	COMBACTE-CARE	IMI	March 2015 – February 2020	Out of scope
Combatting bacterial resistance in Europe - molecules against Gram negative infections	COMBACTE-MAGNET	IMI	January 2015 – December 2021	Out of scope
Collaboration on the optimisation of macromolecular pharmaceutical access	COMPACT	IMI	November 2012 – October 2017	Out of scope

Initiative full name	Initiative acronym	Programme	Approximate start and end dates	Scoring (primary category in bold)
to cellular targets				
Coordinated Research Infrastructures Building Enduring Life-science Services	CORBEL	H2020	September 2015 - August 2019	5
Cancer treatment during pregnancy: from fetal safety to maternal efficacy	CRADLE	Other EU initiative	October 2015 - September 2020	1
Danish Health Data program	Danish Health Data program	National initiative	2014 - 2018	1
Drug Disease Model Resources	DDMoRe	IMI	March 2011 – February 2016	Out of scope
Diabetes research on patient stratification	DIRECT	IMI	February 2012 - January 2019	1
European Infrastructures for translational medicines	EATRIS	Other EU initiative	January 2008 - December 2012	Out of scope
European Bank for induced pluripotent Stem Cells	EbiSC	IMI	January 2014 – December 2016	Out of scope
Communication strategy and tools for optimizing the impact of Ebola vaccination deployment	EBODAC	IMI	December 2014 – November 2017	Out of scope
Ebola and other filoviral haemorrhagic fevers	Ebola+	IMI	February 2015 – January 2017	Out of scope
Ebola virus: modern approaches for developing bedside rapid diagnostics	EbolaMoDRAD	IMI	February 2015 – January 2017	Out of scope
European Clinical Research Infrastructure Network	ECRIN	Other EU initiative	N/A*	Out of scope
eHealth Digital Service Infrastructure	eHDSI	Other EU initiative	2015–2019	Out of scope
Intelligent Knowledge Platform for Personal Health Monitoring Services	eHealthMonitor	FP7	December 2011 - November 2014	Out of scope
Electronic Health Records Systems for Clinical Research	EHR4CR / Insight Platform	IMI	March 2011 - February 2015	5, 1
eHealth in Rheumatology	ELECTOR	H2020	January 2015 – March 2018	Out of scope
European Lead Factory	ELF	IMI	N/A	Out of scope
European LeukemiaNet	ELN	FP6	January 2004 – February 2011	1
European Medical Information Framework	EMIF	IMI	January 2013 - December 2017	1, 5
EMpowering PAtients for a BeTTer Information and improvement of the Communication Systems	EMPATTICS	H2020	N/A	Out of scope
European Medicines Research Training Network	EMTRAIN	IMI	October 2009 – September 2016	Out of scope

Initiative full name	Initiative acronym	Programme	Approximate start and end dates	Scoring (primary category in bold)
European Gram-negative Antibacterial Engine	ENABLE	IMI	February 2014 – January 2020	Out of scope
European Network for Cancer research in Children and Adolescents	ENCCA	FP7	January 2011 - December 2015	1
European Network of Centres for Pharmacoepidemiology and Pharmacovigilance	ENcePP	Other EU initiative	2009 – Ongoing	2, 3, 5
European Network for Genetic and Genomic Epidemiology	ENGAGE	FP7	January 2008 - December 2013	1
European Network of Paediatric Research at the European Medicines Agency	Enpr-EMA	Other EU initiative	2009 – N/A	Out of scope
–Environmental Health Risks in European birth Cohorts	ENRIECO	FP7	March 2009 – March 2011	1
European prevention of Alzeimer's dementia consortium	EPAD	FP7	January 2015 – December 2019	1, 3, 5
European Platform for Rare Disease Registries	EPIRARE	Other EU initiative	April 2011 - April 2014	4, 3
Smart Open Services for European Patients	epSOS	Other EU initiative	July 2008 – June 2014	Out of scope
ERA-NET on translational cancer research in Europe	ERA-NET TRANSCAN	H2020	January 2011 – December 2014	Out of scope
ERA-Net on Rare Diseases	E-Rare	FP7	October 2010 – November 2011	Out of scope
eHealth Standards and Profiles in Action for Europe and Beyond	eStandards	H2020	May 2015 – July 2017	Out of scope
Integrating bioinformatics and chemoinformatics approaches for the development of Expert systems allowing the in silico prediction of toxicities	eTOX	IMI	January 2010 – December 2016	Out of scope
Delivering European Translational Information & Knowledge Management Services	eTRIKS	IMI	October 2012 - September 2017	5, 2
European programme in Pharmacovigilance and Pharmacoepidemiology	Eu2P	IMI	N/A	Out of scope
EU-ADR Alliance project	EU-ADR Alliance project	FP7	February 2008 - January 2012	4
European Autism Interventions - a Multicentre Study for Developing New Medications	EU-AIMS	IMI	September 2009 – June 2016	1
Comparative effectiveness research of existing technologies for	EU-CERT-ICD	FP7	October 2013 – September 2018	Out of scope

Initiative full name	Initiative acronym	Programme	Approximate start and end dates	Scoring (primary category in bold)
prevention, diagnosis and treatment of cardiovascular diseases				
European Patients' Academy on Therapeutic Innovation	EUPATI	IMI	February 2012 – January 2017	Out of scope
Pathophysiology and natural course of autoimmune adrenal failure in Europe	EURADRENAL	FP7	April 2008 - March 2012	1
Enabling information re-Use by linking clinical REsearch and Care	EURECA	H2020	February 2012 - July 2015	4
European Consortium for High-Throughput Research in Rare Kidney Diseases	EuRenOmics	FP7	October 2012 – September 2017	Out of scope
Integration of viral genomics with clinical data to predict response to anti-HIV treatment.	EuResist	FP6	January 2006 – June 2008	Out of scope
European Network of HIV/AIDS Cohort Studies to Coordinate at European and International Level Clinical Research on HIV/AIDS	EuroCoord	H2020	January 2011 – December 2015	1, 2, 4
Safety of Medication use in Pregnancy in Relation to Risk of Congenital Malformations	EUROMEDICAT	FP7	March 2011 - March 2015	1
Pathophysiology and Natural Course of Patients with Primary Antibody Deficiencies	EURO-PADnet	FP7	May 2008 - April 2011	1, 5
Understanding chronic pain and improving its treatment	EUROPAIN	IMI	October 2009 – September 2015	Out of scope
European association of poison centres and clinical toxicologists	EAPCCT	Other EU initiative	N/A	Out of scope
Natural course, Pathomechanisms and Novel Treatment Options in Idiopathic Pulmonary Fibrosis	European IPF Network	FP7	2008 - 2011	1
FAIR data	FAIR data	National imitative	N/A	3
Institute for Health Informatics Research	FARR	National imitative	N/A	1, 2, 3, 4, 5
Ultra-fast molecular filovirus diagnostics	FILODIAG	IMI	February 2015 – January 2017	Out of scope
Standardisation and Development of Assays for Assessment of Influenza Vaccine Correlates of Protection	FluCoP	IMI	March 2015 – February 2020	Out of scope
FrailSafe	FrailSafe	H2020	January 2016 - January 2019	Out of scope

Initiative full name	Initiative acronym	Programme	Approximate start and end dates	Scoring (primary category in bold)
Genotype-To-Phenotype Databases: A Holistic Solution	GEN2PHEN	FP7	2008 - 2013	Out of scope
Genomics England	Genomics England	National initiative		1, 3
Incorporating real-life clinical data into drug development	GetReal	IMI	October 2013 - September 2016	2, 5
Healthcare Alliance for Resourceful Medicines Offensive against Neoplasms in HematologY	HARMONY	IMI	January 2017 – December 2021	Out of scope
Health and Social Care Information Centre	Health and Social Care Information Centre	National initiative	N/A	1, 5
Healthcare IT market intelligence, research and standards organization	HIMSS analytics	Other EU initiative	N/A	Out of scope
European Network for Genetic-Epidemiological Studies: building a method to dissect complex genetic traits, using essential hypertension as a disease model	HYPERGENES	FP7	January 2008 – December 2011	Out of scope
Integration and Interpretation of Information for Individualised Healthcare Network (linked to GEN2PHEN)	I4Health network	FP7	2008 - 2013	Out of scope
Inhaled antibiotics in bronchiectasis and cystic fibrosis	iABC	IMI	August 2015 – July 2020	Out of scope
Improving beta-cell function and identification of diagnostic biomarkers For treatment monitoring in diabetes	IMI-DIA	IMI	February 2010 – September 2015	Out of scope
A pan-national collaborative analytics platform for the exploration of population health	iMoHEALTH	H2020	October 2014 - February 2015	5
Intelligent Assessment of Pharmaceuticals in the Environment	iPiE	IMI	January 2015 – December 2018	Out of scope
IT Future of Medicine	ITFoM	FP7	May 2011 - April 2012	Out of scope
Kinetics for Drug Discovery	K4DD	IMI	November 2012 – October 2017	Out of scope
Platform for Effective Collaborative Clinical Care Management	KareShare	H2020	October 2015 - March 2016	Out of scope
Kconnect	Kconnect	H2020	February 2015 - July 2017	4
LeanBigData	LeanBigData	FP7	February 2015 - February 2018	4, 5
Linked2Safety	Linked2Safety	FP7	October 2011 - September 2014	4, 5

Initiative full name	Initiative acronym	Programme	Approximate start and end dates	Scoring (primary category in bold)
Mapping Chronic Non-Communicable Diseases Research Activities and their Impact	MAPPING_NCD	FP7	January 2014 – December 2015	Out of scope
Biomarkers and molecular tumour classification for non-genotoxic carcinogenesis	MARCAR	IMI	October 2010 – June 2015	Out of scope
Model-Driven European Paediatric Digital Repository	MD PAEDIGREE	FP7	March 2013 – May 2017	Out of scope
Providing the right care to the right patient with MyeloDysplastic Syndrome at the right time	MDS-RIGHT	H2020	May 2015 - 2020	Out of scope
Mechanisms of the Development of ALLergy	MEDALL	FP7	December 2010 – May 2015	1
Development of integrative bioinformatics tools and software applications for analysing huge data sets	MedBioinformatics	H2020	May 2015 – April 2018	Out of scope
The MEtabolic Road to DIAstolic Heart Failure	MEDIA	FP7	January 2011 – June 2016	Out of scope
Microsoft Healthvault	Microsoft Healthvault	National initiative	N/A	Out of scope
Exact Mining from In-Exact Data	MinINexact	FP7	April 2011 - March 2016	4
Mechanism-Based Integrated Systems for the Prediction of Drug-Induced Liver Injury	MIP-DILI	IMI	February 2012 – January 2017	Out of scope
Models of Child Health Appraised	MOCHA	H2020	September 2015 - November 2018	1, 2
Mobile Filovirus Nucleic Acid Test	Mofina	IMI	February 2015 – August 2016	Out of scope
Optimising drug safety monitoring to enhance patient safety and achieve better health outcomes	Monitoring medicines	FP7	September 2009 - July 2013	Out of scope
MyHealthAvatar	MyHealthAvatar	FP7	March 2013 - March 2016	5
Neuromics	Neuromics	FP7	October 2012 - September 2017	1
Novel methods leading to new medications in depression and schizophrenia	NEWMEDS	IMI	September 2009 – August 2014	Out of scope
Next Generation Sequencing for Targatted Personalised Therapy of Leukemia	NGS-PTL	FP7	November 2012 - October 2015	Out of scope
Methods for systematic next generation oncology biomarker development	Onco Track	IMI	January 2011 – December 2015	Out of scope

Initiative full name	Initiative acronym	Programme	Approximate start and end dates	Scoring (primary category in bold)
Open Pharmacological Concepts Triple Store	Open PHACTS	IMI	March 2011 – February 2016	Out of scope
Open electronic health records	OpenEHR	National initiative	N/A	4
Open Mining INfrastructure for TExt and Data	OpenMinTeD	FP7	June 2015 – May 2018	4, 5
Oral biopharmaceutics tools	ORBITO	IMI	October 2012 – September 2017	Out of scope
Childhood and Adolescent Cancer Survivor Care and Follow-up Studies	PANCARESURFUP	FP7	February 2011 – January 2017	1
PatientPartner	PatientPartner	FP7	May 2008 - May 2011	3
Personal Health SystemsForesight	PHS	FP7	September 2012 - September 2014	Out of scope
Long-term PHARMacovigilance for Adverse effects in Childhood arthritis focussing on Immune modulatory drugs	PHARMACHILD	FP7	April 2011- September 2014	1
Prediction of cognitive properties of new drug candidates for neurodegenerative diseases in early clinical development	Pharma-Cog	IMI	January 2010 – December 2014	Out of scope
Pharmaceutical Medicine Training Programme	Pharmatrain	IMI	May 2009 – April 2014	Out of scope
Molecular reclassification to find clinically useful biomarkers for Systemic Autoimmune Diseases	PRECISESADS	IMI	February 2014 - January 2019	Out of scope
Models for preclinical evaluation of drug efficacy in common solid tumours	PredectNew	IMI	February 2011 – January 2016	Out of scope
Personalisation of treatment In Cardiovascular disease through next generation sequencing in Adverse Drug Reactions	PREDICTION ADR	FP7	September 2013 - August 2016	Out of scope
Model-based preclinical development of anti-tuberculosis drug combinations	PreDiCT-TB	IMI	May 2012 – April 2017	Out of scope
Psychiatric Ratings using Intermediate Stratified Markers	PRISM	IMI	April 2016 – March 2019	Out of scope
Physical Activity as a Crucial Patient Reported Outcome in COPD	PRO-active	IMI	September 2009 - May 2016	Out of scope
Pharmacoepidemiological research on outcomes of therapeutics by a	PROTECT	IMI	September 2009 - February 2015	2

Initiative full name	Initiative acronym	Programme	Approximate start and end dates	Scoring (primary category in bold)
European consortium				
Quantative imaging in cancer - connecting cellular process with therapy	Quic-Concept	IMI	September 2011 – December 2017	Out of scope
Remote Assessment of Disease and Relapse in Central Nervous System Disorders	RADAR-CNS	IMI	April 2016 – March 2021	1, 5
Development of rapid point-of-care test platforms for infectious diseases	RAPP-ID	IMI	April 2011 – September 2016	Out of scope
Rare-Bestpractices	RARE-BESTPRACTICES	FP7	January 2013 - December 2016	Out of scope
European Platform of Integrated Information Services for Researchers in the Field of Rare Diseases and Orphan Drugs Supporting Team and Project Building	RareDiseasePlatform	FP7	May 2008 - April 2011	1, 5
Promoting implementation of recommendations on policy, information and data for rare diseases	RD-ACTION	Other EU initiative	June 2015 – May 2018	2, 5
Integrated platform connecting registries, biobanks and clinical bioinformatics for rare disease research	RD-CONNECT	FP7	March 2014 – February 2016	4, 3, 5
Roadmap for European Technologies in Hardware and Networking for Big Data	RETHINK big	FP7	March 2014 - February 2016	5
Research Inventory for Child Health in Europe	RICHE	FP7	N/A	Out of scope
Real World Outcomes Across the AD Spectrum	ROADS	IMI	March 2016 - Ongoing	1, 2
European Modular Education and Training Programme in Safety Sciences for Medicines	SafeSciMET	IMI	January 2010 – September 2016	Out of scope
Safer and Faster Evidence-based Translation	SAFE-T	IMI	June 2009 – June 2015	Out of scope
Scalable, Standard based Interoperability Framework for Sustainable Proactive Post Market Safety Studies	SALUS	FP7	February 2012 - April 2015	4
Strengthening Collaboration for Operating Pharmacovigilance in Europe	SCOPE WP4 (Adverse Drug Reactions collection)	Other EU initiative	2013 - 2016	Out of scope
Strengthening Collaboration for Operating	SCOPE WP8 (Lifecycle Pharmacovigilance)	Other EU initiative	2013 - 2016	Out of scope

Initiative full name	Initiative acronym	Programme	Approximate start and end dates	Scoring (primary category in bold)
Pharmacovigilance in Europe				
Semantic Interoperability for Health Network	SemanticHealthNet	FP7	December 2011 – May 2015	4
Semantic Data Platform for Healthcare	SEMCARE	FP7	January 2014 - December 2015	Out of scope
Services and Health for Elderly in Long TERm care	SHELTER	FP7	January 2009 – December 2011	Out of scope
Grid-enabled pan-Atlantic platform for large scale simulations in paediatric cardiology	Sim-e-Child	FP7	January 2010 – June 2012	Out of scope
European Society for Paediatric Oncology	SIOPE	Other EU initiative	N/A	2
SmartPersonalHealth	SmartPersonalHealth	FP7	January 2010 - December 2011	Out of scope
SoBigData	SoBigData	FP7	September 2015 - August 2018	4, 3, 5
Statistical multi-Omics UNDerstanding of Patient Samples	SOUND	H2020	September 2015 – August 2018	Out of scope
Sarcopenia and physical frailty in older people: multi-component treatment strategies	SPRINTT	IMI	July 2014 – June 2019	Out of scope
Stem cells for biological assays of novel drugs and predictive toxicology	STEMBANCC	IMI	October 2012 – March 2018	Out of scope
Surrogate markers for micro- and macro-vascular hard endpoints for innovative diabetes tools	SUMMIT	IMI	November 2009 – October 2015	Out of scope
Platform as vendor-neutral, open health data platform, designed for real-time, transactional health data storage enabling to go from idea to application in one hour	ThinkEHR Platform	H2020	October 2014 – February 2015	4, 5
Supporting Integrated Data Analysis and Servicing of Alternative Testing Methods in Toxicology	ToxBank	Other EU initiative	January 2011 – December 2015	Out of scope
Transcelerate	Transcelerate	Other EU initiative	2012 – Ongoing	Out of scope
TRANSFoRm	TRANSFoRm	FP7	March 2010 - May 2015	4
Molecular basis of the bacterial cell wall permeability	TRANSLOCATION	IMI	January 2013 – December 2017	Out of scope
Unbiased biomarkers for the prediction of respiratory disease outcomes	U-BIOPRED	IMI	October 2009 – September 2015	1
Unrestricted leveraging of targets for research advancement and drug	ULTRA-DD	IMI	March 2015 – February 2020	Out of scope

Initiative full name	Initiative acronym	Programme	Approximate start and end dates	Scoring (primary category in bold)
discovery				
Establishing the value and business model for sustainable eHealth services in Europe	VALUeHEALTH	H2020	2015 - 2020	5
Mobile phone application for Recognising Adverse Drug Reactions	WEB-RADR	IMI	September 2014 - August 2017	Out of scope
Wearable Sensing and Smart Cloud Computing for Integrated Care to COPD Patients with Co-morbidities	WELCOME	FP7	November 2013 - November 2017	Out of scope
Technological platforms based on insect cells, and/or DNA based RNA virus replicons, and/or prokaryotic systems and/or fungal/plant systems to prepare for surge production of vaccines and antibodies in emergency cases of major viral threats to human and animal health	ZAPI	IMI	March 2015 – February 2020	Out of scope

* N/A: No dates publically found or permanent initiatives

1.2. Excel tracking table columns

- 1.2.1: All 171 initiatives which initially matched at least one of the keywords used in the internet search performed between September and December 2016 were included in the first tabulation of the inventory Excel tracking table, which columns are described in the table below. A tick (for "Yes") or a cross (for "No") was included in the different columns for each initiative following the 1st review of their websites during the 'selection' step:

Column name	Definition
Development of relevant primary data sources, datasets?	Whether any new data sources/datasets have been created within the initiative.
Use of relevant secondary data sources, datasets?	The initiative has used any existing data sources/datasets, has created new cohorts e.g. disease cohorts and/or has linked existing data sources/datasets together.
Development of guidance/methodology?	Creation of a new guidance or methodology on the use of RWD for the evaluation and monitoring of medicinal products relevant to regulators, health care professionals or industry.
Development of relevant governance model?	New governance model comes out of the initiative related to e.g. data privacy access, legal aspects on the use of the data, ethical conduct, patient and data protection, or conflict of interests.
Development of relevant analytical model?	New analytical model e.g. any common protocols or common data models, text analysis/mining techniques or data transformation methods have been developed and promote data linkage.
Development of relevant computing infrastructure / data structure?	New (website/cloud based) platforms to access, share and discover data sources/datasets have been created out of the initiative
Relevant for 2nd step?	Based on the completion of the aforementioned columns, is the initiative considered relevant enough to go through to the next step? Has it developed any tools that could be applied when performing RW studies?

- 1.2.2: The 115 initiatives which went through the 'categorisation' step were included in a second tabulation of the Excel tracking table which columns, or 'attributes', are listed in the table below. For each initiative, a 'Yes' or a 'No' was included in the different columns based on the information that could be identified during the 2nd review of the websites.

Category	Columns
Data source	• Development of relevant primary data sources? • Is the data source longitudinal? • Is the data source disease specific? • Is the data source population specific? • Is the data source project specific? • Unique patient identifier (for possible linkage)? • Is the data source accessible?
Methodology	• Is the initiative developing a new methodological guideline/guidance document? • Use of coded terminology?
Governance model	• Is the initiative developing a new governance model? • Is the initiative developing a code of conduct?
Analytical method	• Is the initiative developing a new common protocol? • Is the initiative developing a new distributed data approach? • Is the initiative developing a new pooling of data? • Is the initiative developing a new analytical tool?
Infrastructure	• Is the initiative developing a new cloud based technology? • Is the initiative developing a new framework/platform for data access, discovery, sharing? • Is the initiative developing innovative technologies? • Is the technology/platform accessible?

2. Information related to the Results section of the manuscript

2.1. Five categories illustrated by 5 scoring dashboards

The 5 dashboards (one for each category) allowed a characterisation of the initiatives through scoring of their main attributes to facilitate identification of the outputs potentially most relevant to specific regulatory questions.

Each column of the dashboards below represents one attribute, which scores ranged from 0 to 3 as indicated at the bottom of each dashboard.

Data source category (1)

Data source name	Primary or secondary Datasource	Longitudinal data	Unique identifier	Access to data	Linkage	Paediatric data included	Patient characteristics	Data on Treatments	Final report published	Outputs match objective	Sustainability mechanism described	Type of care
APPROACH	1	2	2	1	1	0	2	2	0	2	0	2
Astrolab	1	2	1	1	1	2	3	3	0	1	0	2
BigO	1	2	1	1	1	2	3	3	0	2	0	1
BTCure	1	2	1	1	1	0	2	2	0	1	0	1
CENTER-TBI	2	2	2	1	2	2	3	3	0	2	2	2
CHICOS	1	2	2	1	1	2	2	1	2	3	0	1
chILD-EU	1	2	2	1	2	2	2	2	0	1	1	2
CRADLE	1	2	2	1	3	2	2	2	0	2	0	3
Danish Health Data program	1	2	2	1	3	2	2	2	2	2	2	3
DIRECT	2	2	2	1	2	1	2	2	0	2	0	1
EHR4CR / Insight Platform	1	2	2	2	3	2	3	3	2	3	1	2
ELN	1	2	2	1	2	2	3	3	2	1	2	2
EMIF	1	2	2	3	3	2	3	2	0	2	2	2
ENCCA	1	2	2	1	3	2	2	2	2	3	1	2
ENGAGE	1	2	2	1	1	1	2	2	2	3	2	1
ENRIECO	1	2	2	1	2	2	2	2	2	1	0	2
EPAD	2	2	2	3	2	0	2	2	0	2	2	1
EU-AIMS	1	2	1	1	1	2	2	2	1	1	1	1
Euradrenal	1	2	2	1	2	0	2	1	2	3	1	3
EuroCoord	1	2	2	1	2	2	2	3	2	3	1	1
EUROMEDICAT	1	2	2	1	2	2	2	2	2	3	0	2
EURO-PADnet	1	2	2	2	2	2	2	1	2	3	0	1
European IPF Network	1	2	2	1	1	0	1	1	2	3	2	1
FARR	2	2	0	2	2	2	3	3	1	3	1	1
Genomics England	2	2	2	3	2	2	3	1	1	3	2	2
Health and Social Care Information Centre (NHS digital)	1	2	1	1	1	2	2	1	1	3	2	2
MEDALL	1	2	2	1	1	2	3	3	2	3	1	1
MOCHA	1	2	1	3	2	2	2	1	1	2	0	2
Neuromics	2	2	2	1	1	1	2	1	1	2	0	2
PanCareSurFup	1	2	1	1	1	2	3	2	1	1	1	2
PHARMACHILD	2	2	2	1	1	2	3	2	2	3	2	3
RADAR-CNS	1	2	1	1	1	1	1	1	1	2	1	3
RareDiseasePlatform	1	2	1	1	1	2	1	1	2	3	2	2
ROADS	1	2	1	1	1	0	1	1	1	2	1	2
U-BIOPRED	2	2	1	1	1	2	3	3	0	1	0	1
	0 = No 1 = Either primary or secondary 2 = Both	0 = No 1 = Unknown 2 = Yes	0 = No 1 = Unknown 2 = Yes	0 = No 1 = Unknown 2 = Data guardian only 3 = Open to collaborators	0 = No 1 = Unknown 2 = EHR but no drug use 3 = EHR and drug use and others	0 = No 1 = Unknown 2 = Yes	0 = No 1 = Unknown 2 = Basic Demographics only 3 = 2 plus medical history	0 = No 1 = Unknown 2 = Specific to disease only 3 = All prescribed treatments	0 = No 1 = ongoing 2 = Yes	0 = No 1 = Unknown 2 = Ongoing 3 = Yes	0 = No 1 = Unknown 2 = Yes	0 = Unknown 1 = Primary 2 = Secondary / Tertiary 3 = Primary, Secondary, Tertiary

Methodology category (2)

Initiative name	Methodology relevant to RWS?	Methodology on RWS protocol design	Methodology on control bias / confounders	Methodology applicable to any disease and population	Methodology applicable to paediatric population	Methodology applicable to rare diseases	Methodology on how to use EHR / Registries	Standard disease or drug coding terminologies (e.g. Meddra, Snomed, ATC)	Outputs match objectives
ADVANCE	2	2	2	0	2	2	2	2	2
APPROACH	2	0	0	0	0	0	0	2	2
Asterix	2	2	2	0	2	2	2	2	2
BigO	2	0	0	0	2	0	0	0	2
ENCEPP	2	2	2	2	2	2	2	2	3
EPAD	2	2	1	0	0	0	1	0	2
eTRIKS	2	0	0	2	2	2	0	3	2
EuroCoord	2	2	1	0	2	0	1	1	1
FARR	2	2	1	2	2	2	2	0	3
GetReal	2	2	2	2	2	2	0	0	3
MOCHA	2	0	0	0	2	2	0	0	2
PROTECT	2	2	2	2	2	2	2	2	3
RD-ACTION	2	0	0	0	2	2	0	2	2
ROADS	2	2	1	0	2	2	2	1	2
SIOPE	2	0	0	0	2	2	0	0	3
	0 = No 1 = Unknown 2 = Yes	0 = No 1 = Unknown 2 = Yes	0 = No 1 = Unknown 2 = Yes	0 = No 1 = Unknown 2 = Yes	0 = No 1 = Unknown 2 = Yes	0 = No 1 = Unknown 2 = Yes	0 = No 1 = Unknown 2 = Yes	0 = None 1 = Unknown 2 = Either 3 = Both	0 = No 1 = Unknown 2 = Ongoing 3 = Yes

Governance model category (3)

Initiative name	New governance model	Governance applicable to RWS	New code of conduct	Data protection consideration	Governance applicable to paediatric population	Outputs match objectives
ADVANCE	2	2	2	2	2	2
ENCEPP	2	2	2	2	2	3
EPAD	2	2	2	2	0	2
EPIRARE	2	2	0	2	2	3
FAIR data	2	2	2	2	2	3
FARR	2	2	2	2	2	3
Genomics England	1	2	1	2	2	2
PatientPartner	2	2	2	2	2	3
RD-CONNECT	2	2	1	2	2	2
SoBigData	2	2	1	2	2	2
	0 = No 1 = Unknown 2 = Yes	0 = No 1 = Unknown 2 = Yes	0 = No 1 = Unknown 2 = Yes	0 = No 1 = Unknown 2 = Yes	0 = No 1 = Unknown 2 = Yes	0 = No 1 = Unknown 2 = Ongoing 3 = Yes

Analytical model category (4)

Initiative name	New Common protocol	New Common Data Model	Use of OMOP* common data model	Other pooling of data approach	Smart Text Analysis (machine learning, natural language programming)	Data transformation	Applicable outside the initiative?	Outputs match objectives
EPIRARE	0	2	0	2	0	0	1	3
EU-ADR Alliance project	2	0	0	0	2	2	2	3
Eureca	0	2	0	2	2	2	3	1
EuroCoord	0	2	0	2	0	1	2	1
FARR	1	1	0	2	0	2	2	3
Kconnect	0	0	0	0	2	2	0	2
LeanBigData	0	0	0	0	2	0	0	2
Linked2Safety	0	2	0	2	2	0	3	3
MinINexact	0	0	0	0	0	2	1	1
OpenEHR	0	2	0	2	2	0	3	3
OpenMinTeD	0	2	1	2	2	2	3	2
RD-CONNECT	0	2	0	2	2	0	3	2
SALUS	0	2	0	2	0	1	3	3
SemanticHealthNet	0	2	0	2	0	2	3	3
SoBigData	0	0	0	0	2	0	0	2
ThinkEHR Platform	0	2	0	2	0	2	1	1
TRANSFoRm	0	2	1	2	2	2	3	1
		*OMOP: Observational Medical Outcomes Partnership						
	0 = No 1 = Unknown 2 = Yes	0 = No 1 = Unknown 2 = Yes	0 = No 1 = Unknown 2 = Yes	0 = No 1 = Unknown 2 = Yes	0 = No 1 = Unknown 2 = Yes	0 = No 1 = Unknown 2 = Yes	0 = No 1 = Unknown 2 = Partially 3 = Yes	0 = No 1 = Unknown 2 = Ongoing 3 = Yes

Infrastructure category (5)

Initiative name	New cloud based technology	New platform for data sharing and storing	New phone application	Outputs match objectives
ADVANCE	0	2	0	2
APPROACH	0	2	0	2
CORBEL	0	2	0	2
EHR4CR / Insight Platform	0	2	0	3
EMIF	0	2	0	3
ENCEPP	0	2	0	3
EPAD	0	2	0	2
eTRIKS	0	2	0	2
EURO-PADnet	0	2	0	3
FARR	0	2	0	3
GetReal	0	2	0	3
Health and Social Care Information Centre (NHS digital)	0	2	1	3
iMoHEALTH	2	2	0	1
LeanBigData	0	2	0	2
Linked2Safety	0	2	0	3
MyHealthAvatar	0	2	2	3
OpenMinTeD	0	2	0	2
RADAR-CNS	0	2	2	2
RareDiseasePlatform	0	2	0	3
RD-ACTION	0	2	0	2
RD-CONNECT	0	2	0	2
RETHINK big	2	2	1	3
SoBigData	0	2	0	2
ThinkEHR Platform	0	2	2	1
VALUeHEALTH	0	2	0	2
	0 = No 1 = Unknown 2 = Yes	0 = No 1 = Unknown 2 = Yes	0 = No 1 = Unknown 2 = Yes	0 = No 1 = Unknown 2 = Ongoing 3 = Yes

BMJ Open

EU-funded initiatives for Real World Evidence: descriptive analysis of their characteristics and relevance for regulatory decision making

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2018-021864.R1
Article Type:	Research
Date Submitted by the Author:	28-Mar-2018
Complete List of Authors:	Plueschke, Kelly; Pharmacovigilance and Epidemiology McGettigan, Patricia; European Medicines Agency, Pharmacovigilance and Epidemiology Department; William Harvey Research Institute, Barts, and The London School of Medicine and Dentistry, Queen Mary University of London Pacurariu, Alexandra; European Medicines Agency, Pharmacovigilance and Epidemiology Department; Dutch Medicines Evaluation Board Kurz, Xavier; European Medicines Agency, Cave, Alison; European Medicines Agency, Pharmacovigilance and Epidemiology Department
Primary Subject Heading:	Public health
Secondary Subject Heading:	Public health, Global health, Health policy, Health services research
Keywords:	EU-funded initiatives, Real World Evidence, data sources, methodologies, governance and analytical models, infrastructure

EU-funded initiatives for Real World Evidence: descriptive analysis of their characteristics and relevance for regulatory decision making

Kelly Plueschke[1], Patricia McGettigan[1, 2], Alexandra Pacurariu[1, 3], Xavier Kurz[1], Alison Cave[1]

[1] Pharmacovigilance and Epidemiology Department, Inspections Human Medicines Pharmacovigilance and Committees Division, European Medicines Agency (EMA), London E14 5EU, United Kingdom

[2] William Harvey Research Institute, Barts, and The London School of Medicine and Dentistry, Queen Mary University of London, London EC1M 6BQ

[3] Dutch Medicines Evaluation Board, 3531 AH Utrecht, The Netherlands

Corresponding author:

Kelly Plueschke, kelly.plueschke@ema.europa.eu, +44 (0)20 3660 8531

Word count:

2987 (excluding title page, abstract, article summary, references, list of figures' and table's titles, funding, conflicts of interest, disclaimer, acknowledgement)

Number of figures and tables:

6 figures and 1 table in the manuscript

5 figures and 3 tables in the supplementary file

Key words:

EU-funded initiatives, Real World Evidence, data sources, methodologies, governance models, analytical models, infrastructure, sustainability

ABSTRACT

Introduction

A review of European Union (EU)-funded initiatives linked to 'Real World Evidence' (RWE) was performed to determine whether their outputs could be used for the generation of real-world data able to support the European Medicines Agency (EMA)'s regulatory decision making on medicines.

Method

The initiatives were identified from publicly available websites. Their topics were categorised into 5 areas: 'Data source', 'Methodology', 'Governance model', 'Analytical model' and 'Infrastructure'. To assess their immediate relevance for medicines evaluation, their therapeutic areas were compared to the products recommended for EU approval in 2016 and those included in the EMA pharmaceutical business pipeline.

Results

Of 171 originally identified EU-funded initiatives, 65 were selected based on their primary and secondary objectives (35 'Data source' initiatives, 15 'Methodology', 10 'Governance model', 17 'Analytical model' and 25 'Infrastructure'). These received 734 million Euros of public funding. At the time of evaluation, the published outputs of the 40 completed initiatives did not always match their original objectives. Overall, publicly-available information was limited, data access not explicit, and initiatives' sustainability unclear. The topics matched 8 of 14 therapeutic areas of the products recommended for approval in 2016, and of the 15 therapeutic areas in the 2017-2019 pharmaceutical business pipeline. Haematology, gastroenterology or cardiovascular system were poorly represented in the initiatives.

Conclusions

This landscape of EU-funded initiatives linked to RWE which started before the 31st December 2016 highlighted that the immediate utilisation of their outputs to support regulatory decision making is limited, often due to insufficient available information and discrepancies between outputs and objectives. Furthermore the restricted sustainability of the initiatives impacts on their downstream utility. Multiple projects focussing on the same therapeutic areas increases the likelihood of duplication of both efforts and resources. These issues contribute to gaps in generating RWE for medicines and diminish returns on the public funds invested.

ARTICLE SUMMARY

Strengths and limitations of this study

- This is the first evaluation of EU-funded initiatives linked to 'Real World Evidence' (RWE) that looks at the potential for their outputs to support regulatory decision making on medicines;
- The analysis is based on review of the publicly available information provided by each initiative on their English language websites ;
- The internet search was based on a list of internally agreed keywords which might not be fully exhaustive;
- The follow-up period to perform the 2nd and 3rd steps of our analysis was 6 months (January to June 2017). This might not have been long enough to cover the time lag between the finalisation of some of the initiatives and the publication of their final reports, and therefore our analysis may not have taken into account their final outputs;
- The website of each initiative was reviewed by individual EMA staff members. However in some cases there was (i) limited information published on the websites, (ii) broken links and (ii) both

limited access / limited information on access to data, which made it difficult to determine the appropriateness of the initiatives' attributes for inclusion in the inventory, and conclude on their general applicability to regulatory science.

INTRODUCTION

The clinical evidence collected for the marketing authorisation of new medicines traditionally comes from randomised clinical trials (RCTs) but it is recognised that RCT data have limitations including tightly-controlled conditions of clinical care, highly selected populations, and in some scenarios, small sample sizes [1]. As a result, their applicability to the safety and efficacy of medicines in post-authorisation use is unknown. There is therefore a need to supplement RCTs with other sources more representative of everyday 'real world' medical practice in order to provide additional insight on the benefit-risk balance.

According to the GetReal Glossary of Definitions of Common Terms [2], 'Real World Evidence' (RWE) derives from the analysis and/or synthesis of 'Real World Data' (RWD) that can either be primary data collected in a manner which reflects how interventions would be used in routine clinical practice, or secondary data derived from routinely collected data. The range of RWD is wide and sources include electronic healthcare records, patient/disease registries, hospital records and health insurance data/claims databases. The EU Network Strategy to 2020 identifies RWE as a key enabler to bring innovative products to patients with unmet medical needs and to support the safe and effective use of medicines [3].

Whilst the EMA already uses RWE sources in its evaluations, this is typically on an 'ad hoc' basis. There is therefore a need to systematically understand RWE outputs at an EU wide level in order to make best use of existing information and to identify areas where further efforts are needed. We therefore created an inventory of EU-funded initiatives linked to RWE and RWD. The objectives were: i) to identify the initiatives established in terms of RWD sources, relevant methodologies, governance models, analytical models and infrastructures created to facilitate RWD collection, transformation, sharing, and analysis; ii) to understand if and how these initiatives could be exploited to support regulatory decision making on medicines [4]; iii) to consider the initiatives strategically in terms of needs, gaps and opportunities for the generation of RWD related to the therapeutic areas of medicinal products in the pharmaceutical development pipeline.

METHODS

Selection (1st step)

A 3-step approach was followed (Figure 1), starting with an internet search performed by one reviewer (KP) to identify completed and ongoing RWE initiatives funded through the 6th and 7th Framework Programmes (FP6/FP7) [5, 6], Horizon 2020 (H2020) [7] (including the Innovative Medicines Initiative (IMI) [8]), other EU initiatives (e.g. European Research Council, Strengthening Collaboration for Operating Pharmacovigilance in Europe (SCOPE) Joint Action [9], HIMSS EUROPE GOVERNING COUNCIL [10]) and nationally-funded initiatives discovered during the searches. The search terms used were: registry/ies, real world evidence, clinical, electronic medical record, eHealth, big data, electronic health record, data linkage/link, paediatrics, pregnancy, geriatrics, hospital data, data source, unique patient identifier, coding terminology, governance model, common protocol, distributing data approach, pooling of data, and analytical model. The search cut-off date was 31st December 2016; as such all initiatives identified via the search terms which began before or on that date and were funded by either the EU or national bodies were included.

Initiatives started on or before this date that matched any of the keywords were included in an Excel tracking table (Section *1.1.*, *Supplementary file).* Based on their website information, they were categorised according to whether they had developed potentially relevant 1) RWD sources, 2) methodologies, 3) governance models, 4) analytical models or 5) infrastructure that could help provide evidence to support regulatory decision making (Section *1.2.1., Supplementary file)*. Initiatives could fall into more than one category. Those that did not focus on one or more of the categories or that focused only on the early stage of drug development (e.g. pre-clinical or clinical trials, development of molecular compounds/biomarkers), on the '-omics' (e.g. genomics), or on diseases with limited geographical spread (e.g. Ebola outbreak) (Figure 2) were considered out of scope of the inventory and were therefore not selected for further analysis.

Category attributes (2nd step)

For each of the five categories, attributes were defined to facilitate selection of the outputs potentially most relevant for assisting in regulatory evaluations (for example for Data Sources, the attributes included collaborator access, database linkage, unique patent identifier, paediatrics data, and so on). Three reviewers (KP, PM, AP) with experience in pharmacovigilance and regulatory science identified the primary and secondary objectives of each initiative, the attributes, as well as the corresponding deliverables through examination of the initiatives' webpages and through reading of published documents including reports, presentations and publications. They then assigned the initiatives to one or more categories as appropriate *(*Sections *1.1* and *1.2.2., Supplementary file).* In cases of doubt, the three reviewers consulted to achieve consensus.

Scoring (3rd step)

A dashboard listing all the attributes in columns was created for each category. Each initiative was first categorised then characterised through identification and scoring of its attributes. Scores ranged from 0 to 2 for 'Yes / No' questions, for example, whether the data source included paediatric data (where 0='No', 1='Unknown' and 2='Yes'); or from 0 to 3 for more qualitative questions e.g. on the type of care covered (0='Unknown', 1='Primary', 2= 'Secondary/Tertiary' and 3='Primary, Secondary, Tertiary').

Two attributes were scored across all dashboards: 1) the availability of published outputs at the end of each completed project, and 2) whether the published outputs fully addressed the originally stated objectives. The dashboards permitted the creation of figures for each of the categories showing the number of initiatives fulfilling each attribute according to the information provided on their websites.

<Figure 1>

Initiative partners and funding

The numbers of partners named in each initiative and the amount of funding awarded were included in each dashboard.

Therapeutic areas of the initiatives, new medicinal products and EMA business pipeline

The therapeutic areas of the initiatives included in the Data source category were mapped to:

- those of the products recommended for approval through the EU centralised authorisation procedure by the Committee for Medicinal Products for Human Use (CHMP) [11] in 2016;
- those of the products included in the EMA pharmaceutical business pipeline [12] through which developers have expressed a clear intent to submit a marketing authorisation application (MAA) through the centralised procedure to the EMA between March 2017 and December 2019.

Patient involvement statement: This descriptive analysis did not involve any patients.

RESULTS

The first screening returned 171 potentially relevant initiatives (Section *1.1., Supplementary file)* that matched at least one of the search keywords. Following the second screening, 115 initiatives were reviewed for categorisation (51 FP, 30 IMI, 15 H2020, 12 other EU initiatives, 7 nationally funded initiatives), of which 65 fell into one or more of the five categories (Figure 2). Some initiatives were included in more than one category, for example, data source initiatives that also developed infrastructure. The selected initiatives had a median duration of 5 years.

<Figure 2>

Dashboards scoring and summary of findings

Initiatives were scored in the dashboards according to their category(ies) attributes. The main findings for each category are summarised below and illustrated in Figure 3. The scoring dashboards are provided in Section *2.1.* of the *Supplementary file*.

- Data Sources: Information on access to the data sources developed through the initiative was available on the websites of 5 of the 35 Data source initiatives, including 2 where only the data guardian had access, and 3 where access was open to collaborators. Access information was either not publically available or unclear for the remaining 30 initiatives.
- Methodologies: Fifteen initiatives developed methodologies that could be applied to RWD studies, including guidance on protocol design (9), on the management of bias / confounders (5), and on the use of electronic health records and/or registry data (6). Seven initiatives referred to the use of established diseases or drug coding terminologies such as MedDRA, Snomed CT, ATC or Orphanet.
- Governance models: Ten initiatives dealt with confidentiality and data protection aspects, while 6 related to a new code of conduct.
- Analytical models: Ten of 17 initiatives aimed to develop smart text analysis tools such as machine learning, natural language programming, or data mining. Data transformation (anonymisation, right-protection and compression) was referred to in 9 initiatives.
- Infrastructure: All 25 initiatives related to the development of platforms or websites to share, extract and store data, of which 2 also mentioned cloud based technologies. Three developed a smartphone application that could be used by stakeholders including patients to record personal health information or report suspected adverse drug reactions.

<Figure 3>

Outputs versus objectives

The outputs matched the objectives for all 5 of the completed 'Governance model' initiatives (100%), 5 of the 6 completed 'Methodology' initiatives (83%), 10 of the 12 completed 'Infrastructure' initiatives (83%), 15 of the 23 completed 'Data source' initiatives (65%) and for 7 of the 12 completed 'Analytical model' initiatives (58%) (Table 1).

Table 1: Number of completed initiatives with outputs matching objectives

Categories (number of initiatives)	Initiatives Completed / Ongoing	Output match objectives	Unclear if Output match objectives
Data source (n= 35)	23 / 12	15 (65%)	8 (35%)

Categories (number of initiatives)	Initiatives Completed / Ongoing	Output match objectives	Unclear if Output match objectives
Methodology (n=15)	6 / 9	5 (83%)	1 (17%)
Governance model (n=10)	5 / 5	5 (100%)	0
Analytical model (n=17)	12 / 5	7 (58%)	5 (42%)
Infrastructure (n=25)	12 / 13	10 (83%)	2 (17%)

Initiatives' sustainability

Whilst some initiatives continued beyond their agreed timelines through new EU-funded programmes within e.g. H2020, or were followed-up by existing networks (e.g. foundations, associations), others ceased at the end of their allocated time and budget without a sustainability plan for any of the outputs delivered. Of the 23 completed 'Data source' initiatives, 7 had a clear sustainability mechanism described on their website, for example the EMIF Project [13]. Some more recent initiatives integrate roadmaps for future development and continuation as part of their deliverables, like RETHINK big [14].

Comparison of therapeutic areas

Of the 14 different therapeutic areas where products were recommended by the EMA for approval in 2016 and of the 15 represented in the EMA business pipeline, 8 were studied by the initiatives (Figures 4, 5). These included oncology (5 initiatives), neurology/central nervous system (CNS) (7), respiratory (5), immunology (5), rheumatology (3) and metabolic diseases (2). Other therapeutic areas where new products were about to be approved or marketing authorisation applications submitted were not represented in the initiatives, including in particular haematology, gastroenterology or cardiovascular system.

<Figures 4 and 5>

Number of partners and funding

Information on funding was available on the websites of 53 of the 65 initiatives categorised and scored. The total funding was 734 million Euros ranging from 1 to 56 million Euros. Information on the numbers of partners was provided for 58 initiatives and ranged from 1 to 161 partners.

For large initiatives (>20 partners), there was no clear relation between the number of partners and the amount of funding (Figure 6, funding of the 30 'Data source' initiatives). The same partners were involved in multiple initiatives, especially pharmaceutical companies as well as academic centres.

<Figure 6>

DISCUSSION

To our knowledge, this is the first evaluation of EU-funded initiatives in terms of their potential to develop RWD sources, methodologies, governance models, analytical models or infrastructure to contribute to the utility of RWE. It highlights discrepancies between the identified initiatives' objectives and their outputs, a lack of sustainability of the outputs arising from the initiatives, and a mismatch with the therapeutic areas of drugs recently recommended for approval by the EMA and appearing in its business pipeline. Nevertheless the inventory can be consulted by the EMA upon receipt of enquiries arising in the course of regulatory assessments of medicinal products; navigation through the 5

dashboards and filtering of the relevant attributes might provide sources of additional evidence to support decision-making.

Observation

Of 35 'Data source' initiatives, 27 developed RWD sources that focused on diseases in 8 specific therapeutic areas, more than the half of which (17) clustered in neurology/CNS, oncology, and respiratory (Figures 4 and 5). These multiple parallel initiatives with substantial number of partners sometimes involved in the same projects may increase the risk of 1) duplication of efforts and resources, 2) overlap of final outputs, and 3) generating unsustainable and non-interoperable data sources. Most initiatives consisted of standalone projects and while a handful provided a clear set of definitions, most provided no apparent plan to enable the delivery of a sustainable platform on e.g. future data generation, access and reuse. This may result from an omission of the funding bodies to request such a plan or from the lack of incentives e.g. the lack of ongoing funding options.

The inventory highlights that some areas of new product development such as haematology, gastroenterology and cardiovascular system are not represented in the RWD sources identified in the inventory. Therefore opportunities to generate RWD to support evaluation and decision-making on new products could be created through new initiatives targeting these therapeutic areas. It is possible that such omissions will be addressed to some extent by the strategic research agenda of the recently-launched IMI 2 [15] (e.g. Big Data for Better Outcomes Programme [16]).

Whilst the inventory focuses on EU funded initiatives, the same exercise could be extended to international initiatives in order to enlarge the geographical spread and number of data sources, provide a better understanding of populations studied in real world settings worldwide and widen the possibility of linkage between data sources through new analytical models.

Limitations

The three consecutive steps were based on the review of the English version of the websites by individual EMA staff members. The internet search was based on an agreed list of keywords which might not have been fully exhaustive and therefore some relevant initiatives may be missing. The follow-up period to perform the 2nd and 3rd steps of our analysis was only 6 months (January to June 2017), and therefore does not take into account any final outcomes of the originally identified initiatives that were published afterwards. However in general, limited information published on the websites (e.g. lack of final reports or accounts of deliverables while some initiatives did not even have a website), broken links and both limited access / limited information on access to the data, made it sometimes difficult to determine the initiatives' attributes for inclusion in the inventory, let alone conclude on their general applicability to regulatory science. As the initiatives were publicly funded, it is reasonable to expect that their websites would provide up to date information including published outputs as well as some indication in relation to mechanisms of data access and sustainability. Adoption of the suggestion of Galsworthy et al. to develop a central EU repository to make outputs permanently accessible for open meta-analysis and data reuse would go some way to improving this deficit [17, 18]. For initiatives whose data contributes to post-authorisation studies, their registration in the publicly accessible EU PAS Register® provides an opportunity to increase the dissemination of methods and results [19].**Proposed way forward**

From our review, there is an apparent lack of an overarching vision for an infrastructure that would provide sustainable access to the data generated by the projects. Having said that, most recent initiatives include roadmaps as part of their deliverables which may potentially assist in delivering sustainability in the future. Moreover the 11th Call for Proposals launched by the IMI 2 Joint Undertaking [20] aims to provide solutions to ensure that significant results from IMI projects become

fully exploitable, available to all relevant end users, and/or fully sustainable in the long term and in their own right.

Based on our analysis, we highlight the following additional options that stakeholders including the European Commission, regulators, academia, sponsors and advocacy groups may wish to consider for ensuring better exploitation of funded initiatives:

- Regulators and relevant end-users of the final deliverables should be involved early on in the planning and definition of the EU-funded initiatives.
- To streamline efforts, resources, and promote interoperability between outputs, consideration must be given to existing data and lessons learnt from past projects. This should help the identification of the real public health needs to be addressed by new initiatives.
- If timescales allow, future EU RWE initiatives could take account of the EMA business pipeline on medicinal products to avoid gaps in the generation of RWD and ensure regulatory needs are supported.
- Consideration could be given to the development of an agreed set of common data elements across data sources to promote harmonisation and interoperability, especially across those focussed on the same therapeutic areas.
- Use of real world data to support regulatory decision making requires an in depth understanding of the capture, characteristics, quality and validity of the data. Better quantification and information on these areas would enhance utility in the regulatory setting, e.g. through the use of the available EMA procedure on qualification of novel methodologies for medicine development [21].
- The maintenance of the initiatives' websites at the end of the funding period is key to ensure stakeholders are able to keep up to date on the progress made and on the deliverables/achievements. Maintenance of these websites should permit easy public access to information, related reports and peer reviewed publications but would require dedicated funding.
- There is a need for clarity on the possibilities to access RWD sources generated by the initiatives to allow their reuse in other contexts.
- Consideration of explicit mechanisms to ensure the sustainability of outputs delivered by initiatives should be a priority and should be a requirement for any projects' proposals. The availability of funding options to support the ongoing maintenance of RWD resources would incentivise their development and ensure as wide an access as possible.

CONCLUSION

The development of the inventory assists in understanding the extent of existing RWD resources emerging from EU funded initiatives and highlights multiple shortcomings. Ideally results of such initiatives should be reused and sustained and any lessons learnt disseminated. However despite the potential of some initiatives to provide RWE that would support decision-making on the safety of medicinal products, there are challenges for their utilisation in a regulatory context due to the obstacles in the exploitation of their outputs (e.g. limited data access, lack of sustainability).

Gaps and opportunities were identified in terms of specific therapeutic areas which may require RWD but were not a focus of funded initiatives. A number of solutions are proposed to enable better streamlining, communication and sustainability of the outputs generated through the EU funded initiatives. As the 65 initiatives together were granted 734 million Euros of public funding, it is imperative that the shortcomings highlighted here are addressed in future funding programs. This

would go some way to ensuring delivery of stated objectives, data availability, sustainability and reflection of areas of medical need.

REFERENCES

1. Vandenbroucke, J.P., *Observational research, randomised trials, and two views of medical science.* PLoS Med, 2008. **5**(3): p. e67.
2. GetReal. *GetReal WP1 - Deliverable D1.3 "Glossary of Definitions of Common Terms"*. 2016; Available from: http://www.imi-getreal.eu/Portals/1/Documents/01%20deliverables/D1.3%20-%20Revised%20GetReal%20glossary%20-%20FINAL%20updated%20version_25Oct16_webversion.pdf.
3. *EU Medicines Agencies Network Strategy to 2020 - Working together to improve health*. December 2015; Available from: http://www.ema.europa.eu/docs/en_GB/document_library/Other/2015/12/WC500199060.pdf.
4. Goldman, M., N. Seigneuret, and H.G. Eichler, *The Innovative Medicines Initiative: an engine for regulatory science.* Nat Rev Drug Discov, 2015. **14**(1): p. 1-2.
5. Commission, E. *6th Framework Programmes (FP6)*. 2002-2006; Available from: https://ec.europa.eu/research/fp6/index_en.cfm.
6. Commission, E. *7th Framework Programmes (FP7)*. 2014-2020; Available from: https://ec.europa.eu/research/fp7/index_en.cfm.
7. Commission, E. *Horizon 2020 (H2020)*. 2018-2020; Available from: http://ec.europa.eu/programmes/horizon2020/.
8. *Innovative Medicines Initiative*. Available from: https://www.imi.europa.eu/.
9. SCOPE. *European Research Council, Strengthening Collaboration for Operating Pharmacovigilance in Europe* Available from: http://www.scopejointaction.eu/.
10. *HIMSS EUROPE GOVERNING COUNCIL*. Available from: http://www.himss.eu/about-himss/governing-council.
11. *Committee for Medicinal Products for Human Use (CHMP)*. Available from: http://www.ema.europa.eu/ema/index.jsp?curl=pages/about_us/general/general_content_000094.jsp&mid=WC0b01ac0580028c79.
12. *EMA publication on Business pipeline*. 2015; Available from: http://www.ema.europa.eu/docs/en_GB/document_library/Leaflet/2011/03/WC500104229.pdf .
13. *The EMIF Project*. Available from: http://www.emif.eu/.
14. *RETHINK Big*. Available from: http://www.rethinkbig-project.eu/.
15. *Strategic Research Agenda for Innovative Medicines Initiative 2*. Available from: http://www.imi.europa.eu/sites/default/files/uploads/documents/IMI2_SRA_March2014.pdf.
16. *Big Data for Better Outcomes Programme (BD4BO)*. Available from: http://bd4bo.eu/.
17. Galsworthy, M.J., et al., *Academic output of 9 years of EU investment into health research.* Lancet, 2012. **380**(9846): p. 971-2.
18. Molokhia, M., et al., *Academic output from EU-funded health research projects.* Lancet, 2012. **380**(9857): p. 1903-4.
19. Register, E.P. *The European Union electronic Register of Post-Authorisation Studies*. Available from: http://www.encepp.eu/encepp_studies/indexRegister.shtml.
20. *Innovative Medicines Initiatives 2 Joint Understanding 11th Call for Proposals on Exploitation of IMI project results*. Available from: https://www.imi.europa.eu/sites/default/files/uploads/documents/IMI2Call11/IMI2_Call11_CallText.pdf.
21. *Qualification of novel methodologies for medicine development*. Available from: http://www.ema.europa.eu/ema/index.jsp?curl=pages/regulation/document_listing/document_listing_000319.jsp&mid=WC0b01ac0580022bb0

LIST OF FIGURES' AND TABLE'S TITLES

Figure 5 - Specific therapeutic areas of products included in the EMA business pipeline (March 2017 to December 2019) versus those covered by the identified initiatives

Figure 6 - Number of partners versus funding of 'Data source' initiatives

Figure 6 - Numerical data for the number of partners versus funding of Data source initiatives

FOOTNOTES

Acknowledgement: The authors thank Enrico Tognana at the European Medicines Agency for providing data on the EMA pharmaceutical business pipeline covering the period from March 2017 to December 2019, as well as the reviewers for the revision of this manuscript.

Contributions: AC and XK conceived the idea of the descriptive analysis. All authors were involved in the study design. KP, PM and AP performed the analysis and interpretation of results. All authors contributed to writing the manuscript and approved the final draft.

Funding: No sources of funding were used to assist in the preparation of this article.

Conflict of interest: We have read and understood the BMJ Group policy on declaration of interests and declare the following interests: None

Disclaimer: The views expressed in this article are the personal views of the authors and may not be understood or quoted as being made on behalf of or reflecting the position of the European Medicines Agency or one of its committees or working parties.

Ethical approval: Not required.

Data sharing: All the information provided in this manuscript other than the data on the EMA business pipeline is publically available online.

Transparency: The authors affirm that the manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned (and, if relevant, registered) have been explained.

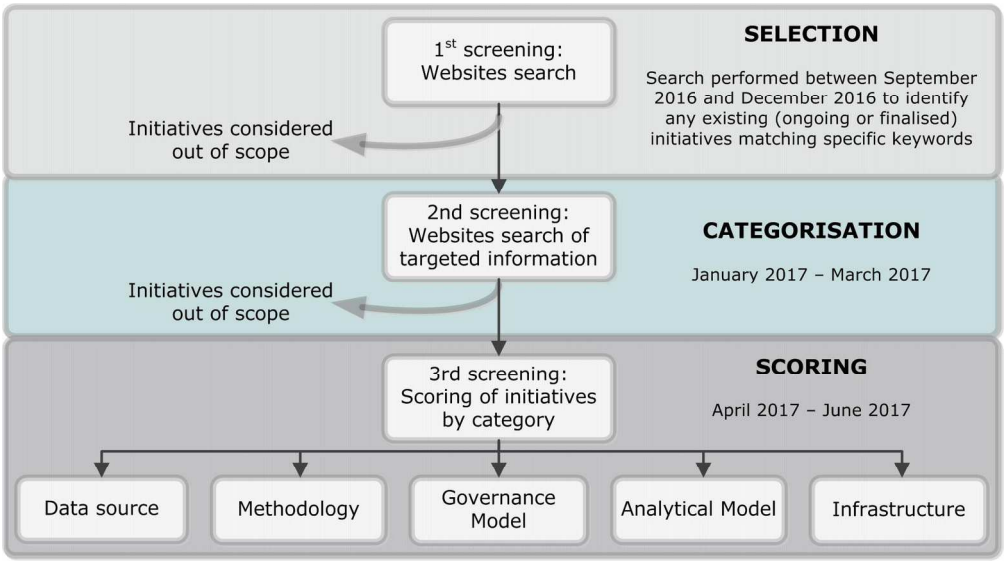

Figure 1 - Development steps of the inventory of EU-funded RWE initiatives

166x92mm (300 x 300 DPI)

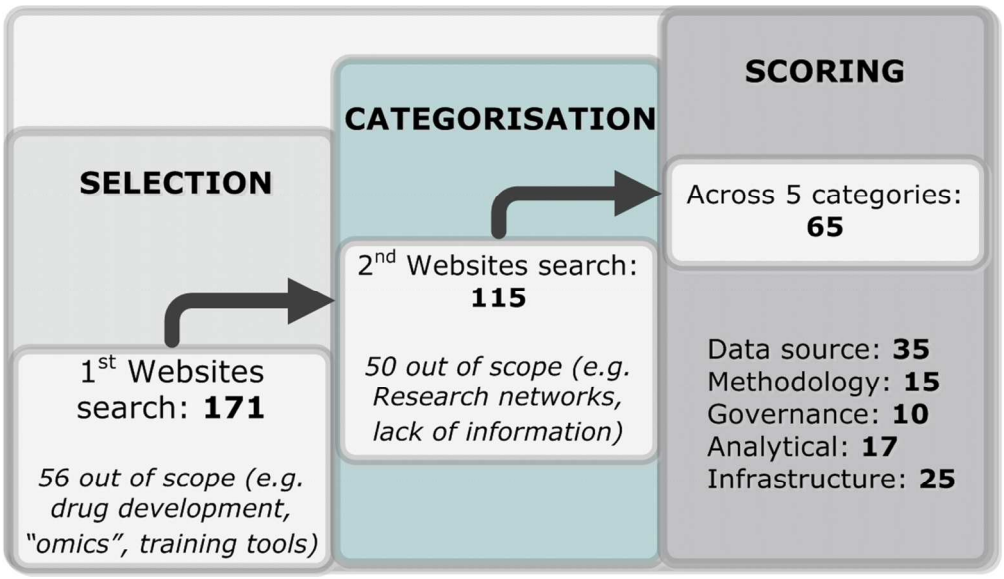

Figure 2 - Selections of initiatives for the inventory of EU-funded RWE initiatives

104x59mm (300 x 300 DPI)

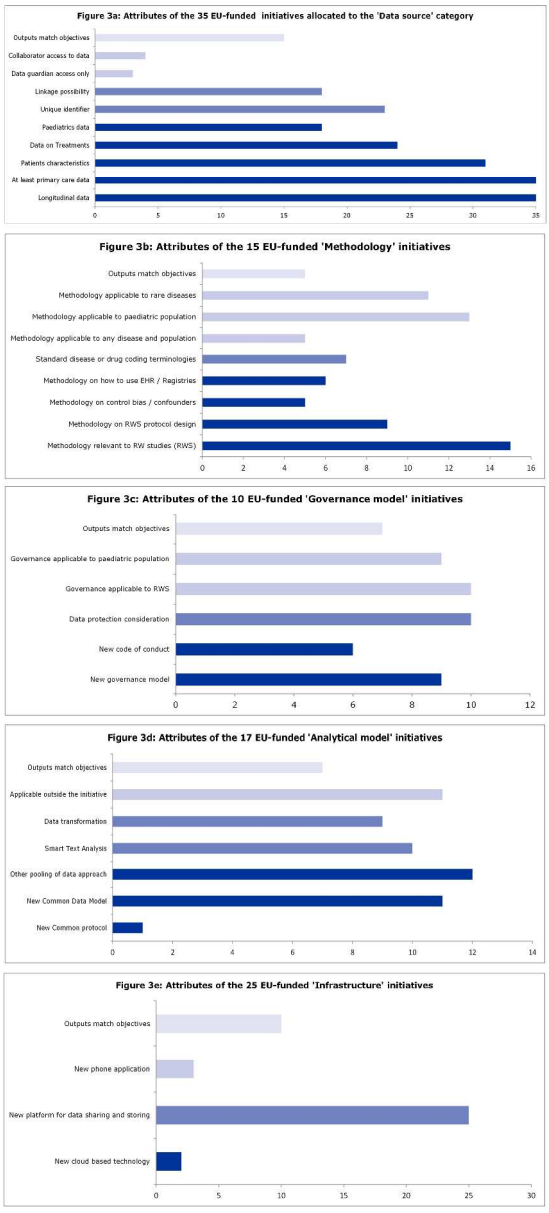

Figure 3 - Attributes of the EU-funded initiatives according to each of the 5 categories

219x484mm (300 x 300 DPI)

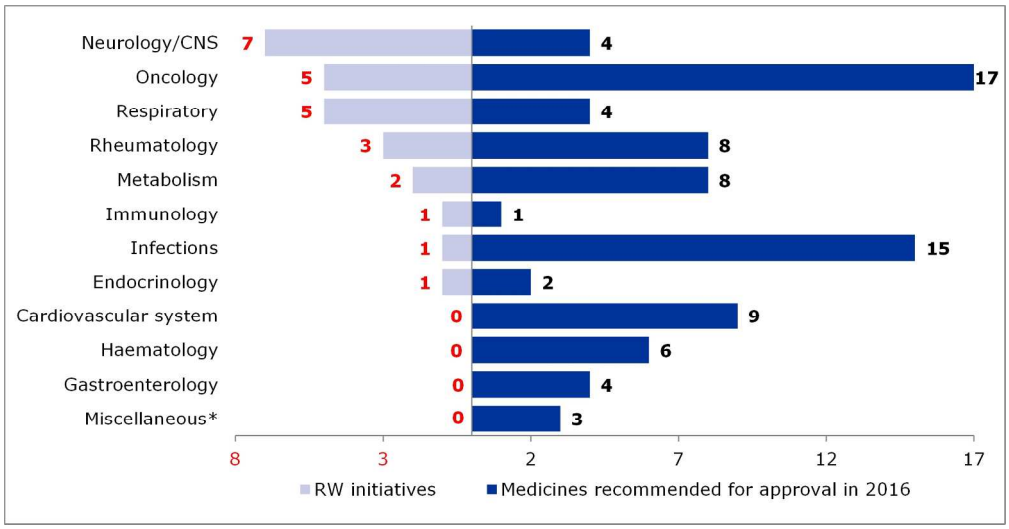

Figure 4 - Specific therapeutic areas of products recommended for approval in 2016 versus those covered by the identified initiatives

198x102mm (300 x 300 DPI)

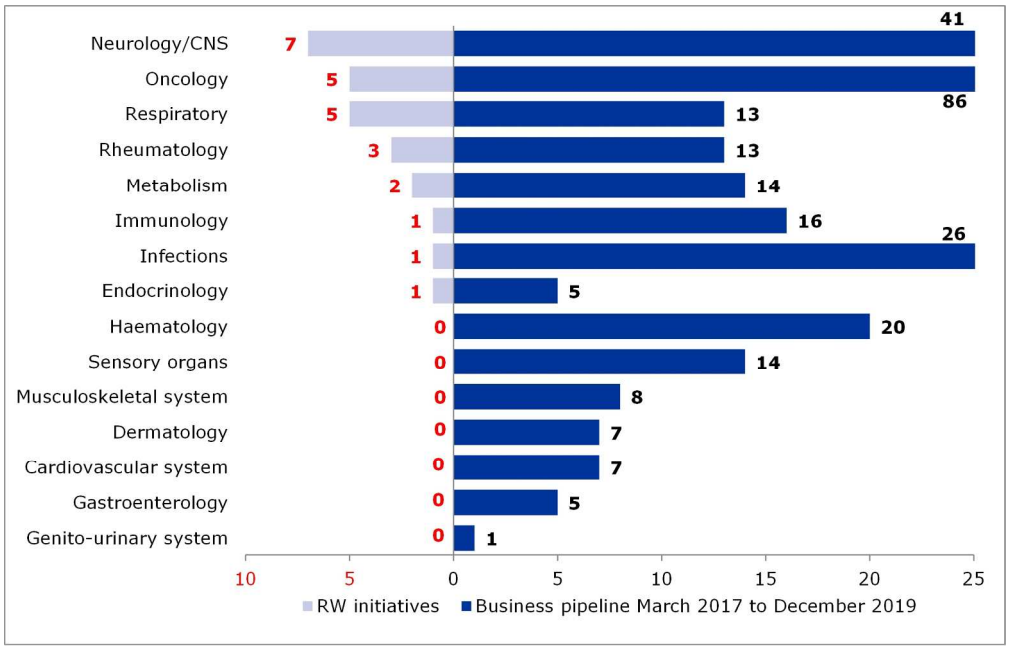

Figure 5 - Specific therapeutic areas of products included in the EMA business pipeline (March 2017 to December 2019) versus those covered by the identified initiatives

196x127mm (300 x 300 DPI)

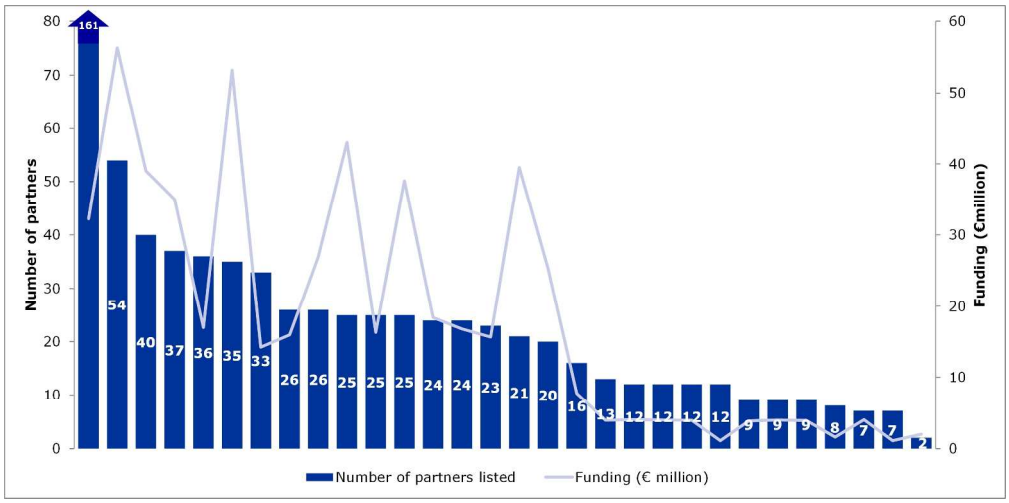

Figure 6 - Number of partners versus funding of Data source initiatives

252x124mm (300 x 300 DPI)

Supplementary file of the manuscript '*EU-funded initiatives for Real World Evidence: descriptive analysis of their characteristics and relevance for regulatory decision making'*

1. Information related to the Methodology section of the manuscript

1.1. List of EU-funded initiatives

The 171 initiatives which initially matched at least one of the keywords used in the internet search performed between September and December 2016 are listed in the table below, together with their EU-funding programme and a link to relevant websites or publications used during the analysis. Some of the initiatives were considered 'out of scope' of the inventory following the 'selection' step (56) or the 'categorisation' step (50). The remaining initiatives (65) underwent 'scoring' across the 5 categories, namely Category 1 'Data source', Category 2 'Methodology', Category 3 'Governance model', Category 4 'Analytical method' and Category 5 'Infrastructure'.

Initiative full name	Initiative acronym	Programme	Approximate start and end dates	Scoring (primary category in bold)
Anti-Biopharmaceutical Immunisation: Prediction and Analysis of Clinical Relevance to Minimise the Risk	ABIRISK	IMI	March 2012 – February 2017	Out of scope
Accelerated Development of Appropriate Patient Therapies -a Sustainable, Multi-stakeholder Approach from Research to Treatment-outcomes	ADAPT SMART	IMI	July 2015 – December 2017	Out of scope
Accelerated development of vaccine benefit-risk collaboration in Europe	ADVANCE	IMI	October 2013 - October 2018	3, 2, 5
Organising mechanistic knowledge about neurodegenerative diseases for the improvement of drug development and therapy	AETIONOMY	IMI	January 2014 - December 2019	Out of scope
Apoptosis systems biology applied to cancer and AIDS	APO-SYS	FP7	February 2008 – January 2012	Out of scope
Applied public-private research enabling osteoarthritis clinical headway	APPROACH	IMI	June 2015 – November 2020	1, 2, 5
Assessing SNOMED CT for Large Scale eHealth Deployments in the EU	ASSESS CT	H2020	February 2015 – December 2016	Out of scope
Advances in Small Trials dEsign for Regulatory Innovation	Asterix	FP7	January 2013 - September 2017	2

Initiative full name	Initiative acronym	Programme	Approximate start and end dates	Scoring (primary category in bold)
and eXcellence				
Assessment of the Safety of Labas in Asthma in Routine Care	Astrolab	FP7	January 2011 - May 2016	1
Big Data EEG Recording and Analysis platform	Big Data EEG – "WhiteBox EEG"	H2020	July 2016 – December 2016	Out of scope
Big data against childhood Obesity	BigO	H2020	January 2016 - November 2020	1, 2
Big Medical Data Use in Primary	BIMEDA	H2020	August 2015 - July 2017	Out of scope
BiobankCloud	BiobankCloud	FP7	December 2012 – November 2015	Out of scope
Building data bridges from biology to medicine in Europe	BIOMEDBRIDGES	FP7	January 2012 – December 2015	Out of scope
Biobank Standardisation and Harmonisation for Research Excellence in the European Union	BioSHaRE-EU	FP7	December 2010 – November 2015	Out of scope
Biomarkers for Enhanced Vaccine Immunosafety	BioVacSafe	IMI	March 2012 - February 2017	Out of scope
Epigenetic blueprint of haematopoietic cells	BLUEPRINT	FP7	October 2011 – September 2016	Out of scope
Be the Cure	BTCure	IMI	April 2011 - March 2017	1
C3-Cloud	C3-Cloud	H2020	May 2016 – April 2020	Out of scope
Collaborative European NeuroTrauma Effectiveness Research in Traumatic Brain Injury	CENTER-TBI	FP7	October 2013 - April 2020	1
Chemical manufacturing methods for the 21st century pharmaceutical industries	CHEM21	IMI	October 2012 – June 2017	Out of scope
Developing a Child Cohort Research Strategy for Europe	CHICOS	FP7	January 2010 - February 2013	1
European Management Platform for Childhood Interstitial Lung Diseases	chILD-EU	FP7	December 2012 - June 2016	1
Combatting Bacterial Resistance in Europe	COMBACTE	IMI	January 2013 – December 2019	Out of scope
Combatting Bacterial Resistance in Europe - Carbapenem Resistance	COMBACTE-CARE	IMI	March 2015 – February 2020	Out of scope
Combatting bacterial resistance in Europe - molecules against Gram negative infections	COMBACTE-MAGNET	IMI	January 2015 – December 2021	Out of scope
Collaboration on the optimisation of macromolecular pharmaceutical access	COMPACT	IMI	November 2012 – October 2017	Out of scope

Initiative full name	Initiative acronym	Programme	Approximate start and end dates	Scoring (primary category in bold)
to cellular targets				
Coordinated Research Infrastructures Building Enduring Life-science Services	CORBEL	H2020	September 2015 - August 2019	5
Cancer treatment during pregnancy: from fetal safety to maternal efficacy	CRADLE	Other EU initiative	October 2015 - September 2020	1
Danish Health Data program	Danish Health Data program	National initiative	2014 - 2018	1
Drug Disease Model Resources	DDMoRe	IMI	March 2011 – February 2016	Out of scope
Diabetes research on patient stratification	DIRECT	IMI	February 2012 - January 2019	1
European Infrastructures for translational medicines	EATRIS	Other EU initiative	January 2008 - December 2012	Out of scope
European Bank for induced pluripotent Stem Cells	EbiSC	IMI	January 2014 – December 2016	Out of scope
Communication strategy and tools for optimizing the impact of Ebola vaccination deployment	EBODAC	IMI	December 2014 – November 2017	Out of scope
Ebola and other filoviral haemorrhagic fevers	Ebola+	IMI	February 2015 – January 2017	Out of scope
Ebola virus: modern approaches for developing bedside rapid diagnostics	EbolaMoDRAD	IMI	February 2015 – January 2017	Out of scope
European Clinical Research Infrastructure Network	ECRIN	Other EU initiative	N/A*	Out of scope
eHealth Digital Service Infrastructure	eHDSI	Other EU initiative	2015–2019	Out of scope
Intelligent Knowledge Platform for Personal Health Monitoring Services	eHealthMonitor	FP7	December 2011 - November 2014	Out of scope
Electronic Health Records Systems for Clinical Research	EHR4CR / Insight Platform	IMI	March 2011 - February 2015	5, 1
eHealth in Rheumatology	ELECTOR	H2020	January 2015 – March 2018	Out of scope
European Lead Factory	ELF	IMI	N/A	Out of scope
European LeukemiaNet	ELN	FP6	January 2004 – February 2011	1
European Medical Information Framework	EMIF	IMI	January 2013 - December 2017	1, 5
EMpowering PAtients for a BeTTer Information and improvement of the Communication Systems	EMPATTICS	H2020	N/A	Out of scope
European Medicines Research Training Network	EMTRAIN	IMI	October 2009 – September 2016	Out of scope

Initiative full name	Initiative acronym	Programme	Approximate start and end dates	Scoring (primary category in bold)
European Gram-negative Antibacterial Engine	ENABLE	IMI	February 2014 – January 2020	Out of scope
European Network for Cancer research in Children and Adolescents	ENCCA	FP7	January 2011 - December 2015	1
European Network of Centres for Pharmacoepidemiology and Pharmacovigilance	ENcePP	Other EU initiative	2009 – Ongoing	2, 3, 5
European Network for Genetic and Genomic Epidemiology	ENGAGE	FP7	January 2008 - December 2013	1
European Network of Paediatric Research at the European Medicines Agency	Enpr-EMA	Other EU initiative	2009 – N/A	Out of scope
–Environmental Health Risks in European birth Cohorts	ENRIECO	FP7	March 2009 – March 2011	1
European prevention of Alzeimer's dementia consortium	EPAD	FP7	January 2015 – December 2019	1, 3, 5
European Platform for Rare Disease Registries	EPIRARE	Other EU initiative	April 2011 - April 2014	4, 3
Smart Open Services for European Patients	epSOS	Other EU initiative	July 2008 – June 2014	Out of scope
ERA-NET on translational cancer research in Europe	ERA-NET TRANSCAN	H2020	January 2011 – December 2014	Out of scope
ERA-Net on Rare Diseases	E-Rare	FP7	October 2010 – November 2011	Out of scope
eHealth Standards and Profiles in Action for Europe and Beyond	eStandards	H2020	May 2015 – July 2017	Out of scope
Integrating bioinformatics and chemoinformatics approaches for the development of Expert systems allowing the in silico prediction of toxicities	eTOX	IMI	January 2010 – December 2016	Out of scope
Delivering European Translational Information & Knowledge Management Services	eTRIKS	IMI	October 2012 - September 2017	5, 2
European programme in Pharmacovigilance and Pharmacoepidemiology	Eu2P	IMI	N/A	Out of scope
EU-ADR Alliance project	EU-ADR Alliance project	FP7	February 2008 - January 2012	4
European Autism Interventions - a Multicentre Study for Developing New Medications	EU-AIMS	IMI	September 2009 – June 2016	1
Comparative effectiveness research of existing technologies for	EU-CERT-ICD	FP7	October 2013 – September 2018	Out of scope

Initiative full name	Initiative acronym	Programme	Approximate start and end dates	Scoring (primary category in bold)
prevention, diagnosis and treatment of cardiovascular diseases				
European Patients' Academy on Therapeutic Innovation	EUPATI	IMI	February 2012 – January 2017	Out of scope
Pathophysiology and natural course of autoimmune adrenal failure in Europe	EURADRENAL	FP7	April 2008 - March 2012	1
Enabling information re-Use by linking clinical REsearch and Care	EURECA	H2020	February 2012 - July 2015	4
European Consortium for High-Throughput Research in Rare Kidney Diseases	EuRenOmics	FP7	October 2012 – September 2017	Out of scope
Integration of viral genomics with clinical data to predict response to anti-HIV treatment.	EuResist	FP6	January 2006 – June 2008	Out of scope
European Network of HIV/AIDS Cohort Studies to Coordinate at European and International Level Clinical Research on HIV/AIDS	EuroCoord	H2020	January 2011 – December 2015	1, 2, 4
Safety of Medication use in Pregnancy in Relation to Risk of Congenital Malformations	EUROMEDICAT	FP7	March 2011 - March 2015	1
Pathophysiology and Natural Course of Patients with Primary Antibody Deficiencies	EURO-PADnet	FP7	May 2008 - April 2011	1, 5
Understanding chronic pain and improving its treatment	EUROPAIN	IMI	October 2009 – September 2015	Out of scope
European association of poison centres and clinical toxicologists	EAPCCT	Other EU initiative	N/A	Out of scope
Natural course, Pathomechanisms and Novel Treatment Options in Idiopathic Pulmonary Fibrosis	European IPF Network	FP7	2008 - 2011	1
FAIR data	FAIR data	National imitative	N/A	3
Institute for Health Informatics Research	FARR	National imitative	N/A	1, 2, 3, 4, 5
Ultra-fast molecular filovirus diagnostics	FILODIAG	IMI	February 2015 – January 2017	Out of scope
Standardisation and Development of Assays for Assessment of Influenza Vaccine Correlates of Protection	FluCoP	IMI	March 2015 – February 2020	Out of scope
FrailSafe	FrailSafe	H2020	January 2016 - January 2019	Out of scope

Initiative full name	Initiative acronym	Programme	Approximate start and end dates	Scoring (primary category in bold)
Genotype-To-Phenotype Databases: A Holistic Solution	GEN2PHEN	FP7	2008 - 2013	Out of scope
Genomics England	Genomics England	National initiative		1, 3
Incorporating real-life clinical data into drug development	GetReal	IMI	October 2013 - September 2016	2, 5
Healthcare Alliance for Resourceful Medicines Offensive against Neoplasms in HematologY	HARMONY	IMI	January 2017 – December 2021	Out of scope
Health and Social Care Information Centre	Health and Social Care Information Centre	National initiative	N/A	1, 5
Healthcare IT market intelligence, research and standards organization	HIMSS analytics	Other EU initiative	N/A	Out of scope
European Network for Genetic-Epidemiological Studies: building a method to dissect complex genetic traits, using essential hypertension as a disease model	HYPERGENES	FP7	January 2008 – December 2011	Out of scope
Integration and Interpretation of Information for Individualised Healthcare Network (linked to GEN2PHEN)	I4Health network	FP7	2008 - 2013	Out of scope
Inhaled antibiotics in bronchiectasis and cystic fibrosis	iABC	IMI	August 2015 – July 2020	Out of scope
Improving beta-cell function and identification of diagnostic biomarkers For treatment monitoring in diabetes	IMI-DIA	IMI	February 2010 – September 2015	Out of scope
A pan-national collaborative analytics platform for the exploration of population health	iMoHEALTH	H2020	October 2014 - February 2015	5
Intelligent Assessment of Pharmaceuticals in the Environment	iPiE	IMI	January 2015 – December 2018	Out of scope
IT Future of Medicine	ITFoM	FP7	May 2011 - April 2012	Out of scope
Kinetics for Drug Discovery	K4DD	IMI	November 2012 – October 2017	Out of scope
Platform for Effective Collaborative Clinical Care Management	KareShare	H2020	October 2015 - March 2016	Out of scope
Kconnect	Kconnect	H2020	February 2015 - July 2017	4
LeanBigData	LeanBigData	FP7	February 2015 - February 2018	4, 5
Linked2Safety	Linked2Safety	FP7	October 2011 - September 2014	4, 5

Initiative full name	Initiative acronym	Programme	Approximate start and end dates	Scoring (primary category in bold)
Mapping Chronic Non-Communicable Diseases Research Activities and their Impact	MAPPING_NCD	FP7	January 2014 – December 2015	Out of scope
Biomarkers and molecular tumour classification for non-genotoxic carcinogenesis	MARCAR	IMI	October 2010 – June 2015	Out of scope
Model-Driven European Paediatric Digital Repository	MD PAEDIGREE	FP7	March 2013 – May 2017	Out of scope
Providing the right care to the right patient with MyeloDysplastic Syndrome at the right time	MDS-RIGHT	H2020	May 2015 - 2020	Out of scope
Mechanisms of the Development of ALLergy	MEDALL	FP7	December 2010 – May 2015	1
Development of integrative bioinformatics tools and software applications for analysing huge data sets	MedBioinformatics	H2020	May 2015 – April 2018	Out of scope
The MEtabolic Road to DIAstolic Heart Failure	MEDIA	FP7	January 2011 – June 2016	Out of scope
Microsoft Healthvault	Microsoft Healthvault	National initiative	N/A	Out of scope
Exact Mining from In-Exact Data	MinINexact	FP7	April 2011 - March 2016	4
Mechanism-Based Integrated Systems for the Prediction of Drug-Induced Liver Injury	MIP-DILI	IMI	February 2012 – January 2017	Out of scope
Models of Child Health Appraised	MOCHA	H2020	September 2015 - November 2018	1, 2
Mobile Filovirus Nucleic Acid Test	Mofina	IMI	February 2015 – August 2016	Out of scope
Optimising drug safety monitoring to enhance patient safety and achieve better health outcomes	Monitoring medicines	FP7	September 2009 - July 2013	Out of scope
MyHealthAvatar	MyHealthAvatar	FP7	March 2013 - March 2016	5
Neuromics	Neuromics	FP7	October 2012 - September 2017	1
Novel methods leading to new medications in depression and schizophrenia	NEWMEDS	IMI	September 2009 – August 2014	Out of scope
Next Generation Sequencing for Targatted Personalised Therapy of Leukemia	NGS-PTL	FP7	November 2012 - October 2015	Out of scope
Methods for systematic next generation oncology biomarker development	Onco Track	IMI	January 2011 – December 2015	Out of scope

Initiative full name	Initiative acronym	Programme	Approximate start and end dates	Scoring (primary category in bold)
Open Pharmacological Concepts Triple Store	Open PHACTS	IMI	March 2011 – February 2016	Out of scope
Open electronic health records	OpenEHR	National initiative	N/A	4
Open Mining INfrastructure for TExt and Data	OpenMinTeD	FP7	June 2015 – May 2018	4, 5
Oral biopharmaceutics tools	ORBITO	IMI	October 2012 – September 2017	Out of scope
Childhood and Adolescent Cancer Survivor Care and Follow-up Studies	PANCARESURFUP	FP7	February 2011 – January 2017	1
PatientPartner	PatientPartner	FP7	May 2008 - May 2011	3
Personal Health SystemsForesight	PHS	FP7	September 2012 - September 2014	Out of scope
Long-term PHARMacovigilance for Adverse effects in Childhood arthritis focussing on Immune modulatory drugs	PHARMACHILD	FP7	April 2011- September 2014	1
Prediction of cognitive properties of new drug candidates for neurodegenerative diseases in early clinical development	Pharma-Cog	IMI	January 2010 – December 2014	Out of scope
Pharmaceutical Medicine Training Programme	Pharmatrain	IMI	May 2009 – April 2014	Out of scope
Molecular reclassification to find clinically useful biomarkers for Systemic Autoimmune Diseases	PRECISESADS	IMI	February 2014 - January 2019	Out of scope
Models for preclinical evaluation of drug efficacy in common solid tumours	PredectNew	IMI	February 2011 – January 2016	Out of scope
Personalisation of treatment In Cardiovascular disease through next generation sequencing in Adverse Drug Reactions	PREDICTION ADR	FP7	September 2013 - August 2016	Out of scope
Model-based preclinical development of anti-tuberculosis drug combinations	PreDiCT-TB	IMI	May 2012 – April 2017	Out of scope
Psychiatric Ratings using Intermediate Stratified Markers	PRISM	IMI	April 2016 – March 2019	Out of scope
Physical Activity as a Crucial Patient Reported Outcome in COPD	PRO-active	IMI	September 2009 - May 2016	Out of scope
Pharmacoepidemiological research on outcomes of therapeutics by a	PROTECT	IMI	September 2009 - February 2015	2

Initiative full name	Initiative acronym	Programme	Approximate start and end dates	Scoring (primary category in bold)
European consortium				
Quantative imaging in cancer - connecting cellular process with therapy	Quic-Concept	IMI	September 2011 – December 2017	Out of scope
Remote Assessment of Disease and Relapse in Central Nervous System Disorders	RADAR-CNS	IMI	April 2016 – March 2021	1, 5
Development of rapid point-of-care test platforms for infectious diseases	RAPP-ID	IMI	April 2011 – September 2016	Out of scope
Rare-Bestpractices	RARE-BESTPRACTICES	FP7	January 2013 - December 2016	Out of scope
European Platform of Integrated Information Services for Researchers in the Field of Rare Diseases and Orphan Drugs Supporting Team and Project Building	RareDiseasePlatform	FP7	May 2008 - April 2011	1, 5
Promoting implementation of recommendations on policy, information and data for rare diseases	RD-ACTION	Other EU initiative	June 2015 – May 2018	2, 5
Integrated platform connecting registries, biobanks and clinical bioinformatics for rare disease research	RD-CONNECT	FP7	March 2014 – February 2016	4, 3, 5
Roadmap for European Technologies in Hardware and Networking for Big Data	RETHINK big	FP7	March 2014 - February 2016	5
Research Inventory for Child Health in Europe	RICHE	FP7	N/A	Out of scope
Real World Outcomes Across the AD Spectrum	ROADS	IMI	March 2016 - Ongoing	1, 2
European Modular Education and Training Programme in Safety Sciences for Medicines	SafeSciMET	IMI	January 2010 – September 2016	Out of scope
Safer and Faster Evidence-based Translation	SAFE-T	IMI	June 2009 – June 2015	Out of scope
Scalable, Standard based Interoperability Framework for Sustainable Proactive Post Market Safety Studies	SALUS	FP7	February 2012 - April 2015	4
Strengthening Collaboration for Operating Pharmacovigilance in Europe	SCOPE WP4 (Adverse Drug Reactions collection)	Other EU initiative	2013 - 2016	Out of scope
Strengthening Collaboration for Operating	SCOPE WP8 (Lifecycle Pharmacovigilance)	Other EU initiative	2013 - 2016	Out of scope

Initiative full name	Initiative acronym	Programme	Approximate start and end dates	Scoring (primary category in bold)
Pharmacovigilance in Europe				
Semantic Interoperability for Health Network	SemanticHealthNet	FP7	December 2011 – May 2015	4
Semantic Data Platform for Healthcare	SEMCARE	FP7	January 2014 - December 2015	Out of scope
Services and Health for Elderly in Long TERm care	SHELTER	FP7	January 2009 – December 2011	Out of scope
Grid-enabled pan-Atlantic platform for large scale simulations in paediatric cardiology	Sim-e-Child	FP7	January 2010 – June 2012	Out of scope
European Society for Paediatric Oncology	SIOPE	Other EU initiative	N/A	2
SmartPersonalHealth	SmartPersonalHealth	FP7	January 2010 - December 2011	Out of scope
SoBigData	SoBigData	FP7	September 2015 - August 2018	4, 3, 5
Statistical multi-Omics UNDerstanding of Patient Samples	SOUND	H2020	September 2015 – August 2018	Out of scope
Sarcopenia and physical frailty in older people: multi-component treatment strategies	SPRINTT	IMI	July 2014 – June 2019	Out of scope
Stem cells for biological assays of novel drugs and predictive toxicology	STEMBANCC	IMI	October 2012 – March 2018	Out of scope
Surrogate markers for micro- and macro-vascular hard endpoints for innovative diabetes tools	SUMMIT	IMI	November 2009 – October 2015	Out of scope
Platform as vendor-neutral, open health data platform, designed for real-time, transactional health data storage enabling to go from idea to application in one hour	ThinkEHR Platform	H2020	October 2014 – February 2015	4, 5
Supporting Integrated Data Analysis and Servicing of Alternative Testing Methods in Toxicology	ToxBank	Other EU initiative	January 2011 – December 2015	Out of scope
Transcelerate	Transcelerate	Other EU initiative	2012 – Ongoing	Out of scope
TRANSFoRm	TRANSFoRm	FP7	March 2010 - May 2015	4
Molecular basis of the bacterial cell wall permeability	TRANSLOCATION	IMI	January 2013 – December 2017	Out of scope
Unbiased biomarkers for the prediction of respiratory disease outcomes	U-BIOPRED	IMI	October 2009 – September 2015	1
Unrestricted leveraging of targets for research advancement and drug	ULTRA-DD	IMI	March 2015 – February 2020	Out of scope

Initiative full name	Initiative acronym	Programme	Approximate start and end dates	Scoring (primary category in bold)
discovery				
Establishing the value and business model for sustainable eHealth services in Europe	VALUeHEALTH	H2020	2015 - 2020	5
Mobile phone application for Recognising Adverse Drug Reactions	WEB-RADR	IMI	September 2014 - August 2017	Out of scope
Wearable Sensing and Smart Cloud Computing for Integrated Care to COPD Patients with Co-morbidities	WELCOME	FP7	November 2013 - November 2017	Out of scope
Technological platforms based on insect cells, and/or DNA based RNA virus replicons, and/or prokaryotic systems and/or fungal/plant systems to prepare for surge production of vaccines and antibodies in emergency cases of major viral threats to human and animal health	ZAPI	IMI	March 2015 – February 2020	Out of scope

* N/A: No dates publically found or permanent initiatives

1.2. Excel tracking table columns

- 1.2.1: All 171 initiatives which initially matched at least one of the keywords used in the internet search performed between September and December 2016 were included in the first tabulation of the inventory Excel tracking table, which columns are described in the table below. A tick (for "Yes") or a cross (for "No") was included in the different columns for each initiative following the 1st review of their websites during the 'selection' step:

Column name	Definition
Development of relevant primary data sources, datasets?	Whether any new data sources/datasets have been created within the initiative.
Use of relevant secondary data sources, datasets?	The initiative has used any existing data sources/datasets, has created new cohorts e.g. disease cohorts and/or has linked existing data sources/datasets together.
Development of guidance/methodology?	Creation of a new guidance or methodology on the use of RWD for the evaluation and monitoring of medicinal products relevant to regulators, health care professionals or industry.
Development of relevant governance model?	New governance model comes out of the initiative related to e.g. data privacy access, legal aspects on the use of the data, ethical conduct, patient and data protection, or conflict of interests.
Development of relevant analytical model?	New analytical model e.g. any common protocols or common data models, text analysis/mining techniques or data transformation methods have been developed and promote data linkage.
Development of relevant computing infrastructure / data structure?	New (website/cloud based) platforms to access, share and discover data sources/datasets have been created out of the initiative
Relevant for 2nd step?	Based on the completion of the aforementioned columns, is the initiative considered relevant enough to go through to the next step? Has it developed any tools that could be applied when performing RW studies?

- 1.2.2: The 115 initiatives which went through the 'categorisation' step were included in a second tabulation of the Excel tracking table which columns, or 'attributes', are listed in the table below. For each initiative, a 'Yes' or a 'No' was included in the different columns based on the information that could be identified during the 2nd review of the websites.

Category	Columns
Data source	• Development of relevant primary data sources? • Is the data source longitudinal? • Is the data source disease specific? • Is the data source population specific? • Is the data source project specific? • Unique patient identifier (for possible linkage)? • Is the data source accessible?
Methodology	• Is the initiative developing a new methodological guideline/guidance document? • Use of coded terminology?
Governance model	• Is the initiative developing a new governance model? • Is the initiative developing a code of conduct?
Analytical method	• Is the initiative developing a new common protocol? • Is the initiative developing a new distributed data approach? • Is the initiative developing a new pooling of data? • Is the initiative developing a new analytical tool?
Infrastructure	• Is the initiative developing a new cloud based technology? • Is the initiative developing a new framework/platform for data access, discovery, sharing? • Is the initiative developing innovative technologies? • Is the technology/platform accessible?

2. Information related to the Results section of the manuscript

2.1. Five categories illustrated by 5 scoring dashboards

The 5 dashboards (one for each category) allowed a characterisation of the initiatives through scoring of their main attributes to facilitate identification of the outputs potentially most relevant to specific regulatory questions.

Each column of the dashboards below represents one attribute, which scores ranged from 0 to 3 as indicated at the bottom of each dashboard.

Data source category (1)

Data source name	Primary or secondary Datasource	Longitudinal data	Unique identifier	Access to data	Linkage	Paediatric data included	Patient characteristics	Data on Treatments	Final report published	Outputs match objective	Sustainability mechanism described	Type of care
APPROACH	1	2	2	1	1	0	2	2	0	2	0	2
Astrolab	1	2	1	1	1	2	3	3	0	1	0	2
BigO	1	2	1	1	1	2	3	3	0	2	0	1
BTCure	1	2	1	1	1	0	2	2	0	1	0	1
CENTER-TBI	2	2	2	1	2	2	3	3	0	2	2	2
CHICOS	1	2	2	1	1	2	2	1	2	3	0	1
chILD-EU	1	2	2	1	2	2	2	2	0	1	1	2
CRADLE	1	2	2	1	3	2	2	2	0	2	0	3
Danish Health Data program	1	2	2	1	3	2	2	2	2	2	2	3
DIRECT	2	2	2	1	2	1	2	2	0	2	0	1
EHR4CR / Insight Platform	1	2	2	2	3	2	3	3	2	3	1	2
ELN	1	2	2	1	2	2	3	3	2	1	2	2
EMIF	1	2	2	3	3	2	3	2	0	2	2	2
ENCCA	1	2	2	1	3	2	2	2	2	3	1	2
ENGAGE	1	2	2	1	1	1	2	2	2	3	2	1
ENRIECO	1	2	2	1	2	2	2	2	2	1	0	2
EPAD	2	2	2	3	2	0	2	2	0	2	2	1
EU-AIMS	1	2	1	1	1	2	2	2	1	1	1	1
Euradrenal	1	2	2	1	2	0	2	1	2	3	1	3
EuroCoord	1	2	2	1	2	2	2	3	2	3	1	1
EUROMEDICAT	1	2	2	1	2	2	2	2	2	3	0	2
EURO-PADnet	1	2	2	2	2	2	2	1	2	3	0	1
European IPF Network	1	2	2	1	1	0	1	1	2	3	2	1
FARR	2	2	0	2	2	2	3	3	1	3	1	1
Genomics England	2	2	2	3	2	2	3	1	1	3	2	2
Health and Social Care Information Centre (NHS digital)	1	2	1	1	1	2	2	1	1	3	2	2
MEDALL	1	2	2	1	1	2	3	3	2	3	1	1
MOCHA	1	2	1	3	2	2	2	1	1	2	0	2
Neuromics	2	2	2	1	1	1	2	1	1	2	0	2
PanCareSurFup	1	2	1	1	1	2	3	2	1	1	1	2
PHARMACHILD	2	2	2	1	1	2	3	2	2	3	2	3
RADAR-CNS	1	2	1	1	1	1	1	1	1	2	1	3
RareDiseasePlatform	1	2	1	1	1	2	1	1	2	3	2	2
ROADS	1	2	1	1	1	0	1	1	1	2	1	2
U-BIOPRED	2	2	1	1	1	2	3	3	0	1	0	1
	0 = No 1 = Either primary or secondary 2 = Both	0 = No 1 = Unknown 2 = Yes	0 = No 1 = Unknown 2 = Yes	0 = No 1 = Unknown 2 = Data guardian only 3 = Open to collaborators	0 = No 1 = Unknown 2 = EHR but no drug use 3 = EHR and drug use and others	0 = No 1 = Unknown 2 = Yes	0 = No 1 = Unknown 2 = Basic Demographics only 3 = 2 plus medical history	0 = No 1 = Unknown 2 = Specific to disease only 3 = All prescribed treatments	0 = No 1 = ongoing 2 = Yes	0 = No 1 = Unknown 2 = Ongoing 3 = Yes	0 = No 1 = Unknown 2 = Yes	0 = Unknown 1 = Primary 2 = Secondary / Tertiary 3 = Primary, Secondary, Tertiary

Methodology category (2)

Initiative name	Methodology relevant to RWS?	Methodology on RWS protocol design	Methodology on control bias / confounders	Methodology applicable to any disease and population	Methodology applicable to paediatric population	Methodology applicable to rare diseases	Methodology on how to use EHR / Registries	Standard disease or drug coding terminologies (e.g. Meddra, Snomed, ATC)	Outputs match objectives
ADVANCE	2	2	2	0	2	2	2	2	2
APPROACH	2	0	0	0	0	0	0	2	2
Asterix	2	2	2	0	2	2	2	2	2
BigO	2	0	0	0	2	0	0	0	2
ENCEPP	2	2	2	2	2	2	2	2	3
EPAD	2	2	1	0	0	0	1	0	2
eTRIKS	2	0	0	2	2	2	0	3	2
EuroCoord	2	2	1	0	2	0	1	1	1
FARR	2	2	1	2	2	2	2	0	3
GetReal	2	2	2	2	2	2	0	0	3
MOCHA	2	0	0	0	2	2	0	0	2
PROTECT	2	2	2	2	2	2	2	2	3
RD-ACTION	2	0	0	0	2	2	0	2	2
ROADS	2	2	1	0	2	2	2	1	2
SIOPE	2	0	0	0	2	2	0	0	3
	0 = No 1 = Unknown 2 = Yes	0 = No 1 = Unknown 2 = Yes	0 = No 1 = Unknown 2 = Yes	0 = No 1 = Unknown 2 = Yes	0 = No 1 = Unknown 2 = Yes	0 = No 1 = Unknown 2 = Yes	0 = No 1 = Unknown 2 = Yes	0 = None 1 = Unknown 2 = Either 3 = Both	0 = No 1 = Unknown 2 = Ongoing 3 = Yes

Governance model category (3)

Initiative name	New governance model	Governance applicable to RWS	New code of conduct	Data protection consideration	Governance applicable to paediatric population	Outputs match objectives
ADVANCE	2	2	2	2	2	2
ENCEPP	2	2	2	2	2	3
EPAD	2	2	2	2	0	2
EPIRARE	2	2	0	2	2	3
FAIR data	2	2	2	2	2	3
FARR	2	2	2	2	2	3
Genomics England	1	2	1	2	2	2
PatientPartner	2	2	2	2	2	3
RD-CONNECT	2	2	1	2	2	2
SoBigData	2	2	1	2	2	2
	0 = No 1 = Unknown 2 = Yes	0 = No 1 = Unknown 2 = Yes	0 = No 1 = Unknown 2 = Yes	0 = No 1 = Unknown 2 = Yes	0 = No 1 = Unknown 2 = Yes	0 = No 1 = Unknown 2 = Ongoing 3 = Yes

Analytical model category (4)

Initiative name	New Common protocol	New Common Data Model	Use of OMOP* common data model	Other pooling of data approach	Smart Text Analysis (machine learning, natural language programming)	Data transformation	Applicable outside the initiative?	Outputs match objectives
EPIRARE	0	2	0	2	0	0	1	3
EU-ADR Alliance project	2	0	0	0	2	2	2	3
Eureca	0	2	0	2	2	2	3	1
EuroCoord	0	2	0	2	0	1	2	1
FARR	1	1	0	2	0	2	2	3
Kconnect	0	0	0	0	2	2	0	2
LeanBigData	0	0	0	0	2	0	0	2
Linked2Safety	0	2	0	2	2	0	3	3
MinINexact	0	0	0	0	0	2	1	1
OpenEHR	0	2	0	2	2	0	3	3
OpenMinTeD	0	2	1	2	2	2	3	2
RD-CONNECT	0	2	0	2	2	0	3	2
SALUS	0	2	0	2	0	1	3	3
SemanticHealthNet	0	2	0	2	0	2	3	3
SoBigData	0	0	0	0	2	0	0	2
ThinkEHR Platform	0	2	0	2	0	2	1	1
TRANSFoRm	0	2	1	2	2	2	3	1
		*OMOP: Observational Medical Outcomes Partnership						
	0 = No 1 = Unknown 2 = Yes	0 = No 1 = Unknown 2 = Yes	0 = No 1 = Unknown 2 = Yes	0 = No 1 = Unknown 2 = Yes	0 = No 1 = Unknown 2 = Yes	0 = No 1 = Unknown 2 = Yes	0 = No 1 = Unknown 2 = Partially 3 = Yes	0 = No 1 = Unknown 2 = Ongoing 3 = Yes

Infrastructure category (5)

Initiative name	New cloud based technology	New platform for data sharing and storing	New phone application	Outputs match objectives
ADVANCE	0	2	0	2
APPROACH	0	2	0	2
CORBEL	0	2	0	2
EHR4CR / Insight Platform	0	2	0	3
EMIF	0	2	0	3
ENCEPP	0	2	0	3
EPAD	0	2	0	2
eTRIKS	0	2	0	2
EURO-PADnet	0	2	0	3
FARR	0	2	0	3
GetReal	0	2	0	3
Health and Social Care Information Centre (NHS digital)	0	2	1	3
iMoHEALTH	2	2	0	1
LeanBigData	0	2	0	2
Linked2Safety	0	2	0	3
MyHealthAvatar	0	2	2	3
OpenMinTeD	0	2	0	2
RADAR-CNS	0	2	2	2
RareDiseasePlatform	0	2	0	3
RD-ACTION	0	2	0	2
RD-CONNECT	0	2	0	2
RETHINK big	2	2	1	3
SoBigData	0	2	0	2
ThinkEHR Platform	0	2	2	1
VALUeHEALTH	0	2	0	2
	0 = No 1 = Unknown 2 = Yes	0 = No 1 = Unknown 2 = Yes	0 = No 1 = Unknown 2 = Yes	0 = No 1 = Unknown 2 = Ongoing 3 = Yes